

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

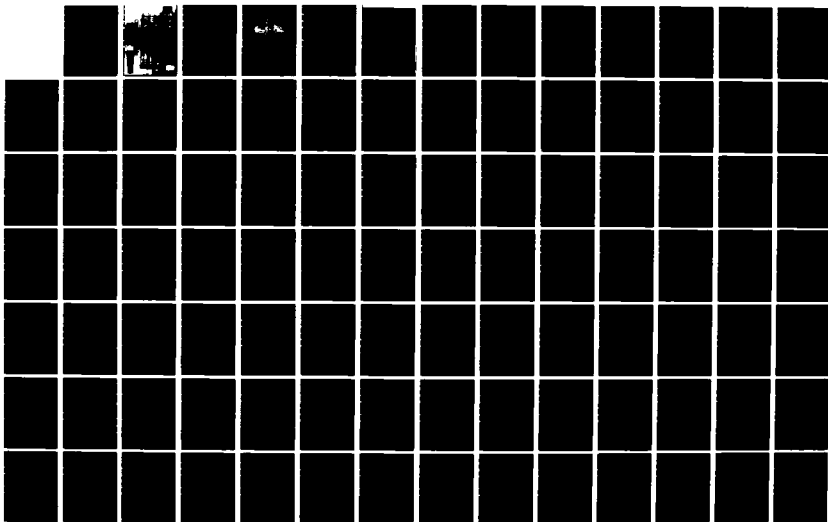
1/13

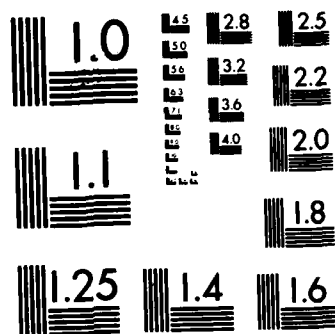
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A



2

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

UNITED STATES AIR FORCE

SUMMER FACULTY  
RESEARCH PROGRAM

conducted by the  
SOUTHEASTERN CENTER  
FOR ELECTRICAL ENGINEERING EDUCATION  
(SCEEE)

1984  
TECHNICAL REPORT  
VOLUME II OF III

WARREN D. PEELE

EARL L. STEELE

PROGRAM DIRECTORS, SCEEE

DISTRIBUTION STATEMENT A

Approved for public release;  
Distribution Unlimited

THE SCEEE PR

85 5 28 158

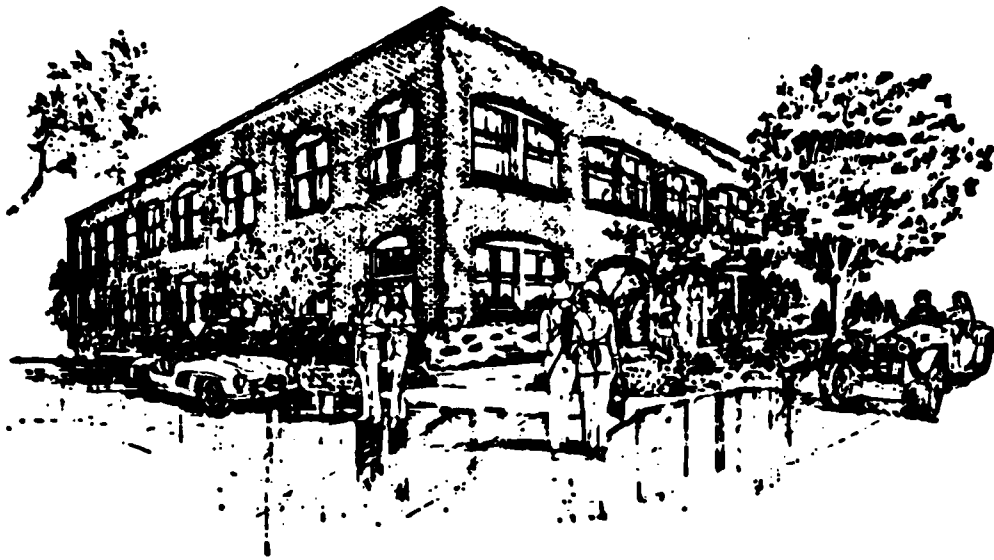
AD-A154 337

DTIC FILE COPY

DTIC  
ELECTE  
MAY 30 1985  
D

AFOSR-TR- 85 - 0480

Approved for public release;  
distribution unlimited.



SCEEE  
©  
1984

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD A 154 337

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b. RESTRICTIVE MARKINGS																						
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT <b>APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED</b>																						
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE																									
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR- 85 - 0480</b>																						
6a. NAME OF PERFORMING ORGANIZATION <b>Southeastern Center for Electrical Engineering Education</b>		6b. OFFICE SYMBOL (If applicable) <b>XOT</b>		7a. NAME OF MONITORING ORGANIZATION <b>Air Force Office of Scientific Research/XOT</b>																					
6c. ADDRESS (City, State and ZIP Code) <b>11th &amp; Massachusetts Ave. St. Cloud, FL 32769</b>		7b. ADDRESS (City, State and ZIP Code) <b>Building 410 Bolling AFB, DC 20332-6448</b>																							
8a. NAME OF FUNDING/SPONSORING ORGANIZATION <b>AFOSR</b>		8b. OFFICE SYMBOL (If applicable) <b>XOT</b>		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <b>F49620-82-C-0035</b>																					
8c. ADDRESS (City, State and ZIP Code) <b>Building 410 Bolling AFB, DC 20332</b>		10. SOURCE OF FUNDING NOS. <table border="1"><tr><td>PROGRAM ELEMENT NO.</td><td>PROJECT NO.</td><td>TASK NO.</td><td>WORK UNIT NO.</td></tr><tr><td>61102F</td><td>2301</td><td>D5</td><td></td></tr></table>				PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.	61102F	2301	D5													
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.																						
61102F	2301	D5																							
11. TITLE (Include Security Classification) <b>United States Air Force Summer Faculty Research Program (1984) Volume 2</b>																									
12. PERSONAL AUTHOR(S) <b>Warren D. Peele, Earl L. Steele, Major Amos L. Otis</b>																									
13a. TYPE OF REPORT <b>Final</b>		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Yr., Mo., Day) <b>December 1984</b>																					
15. PAGE COUNT <b>15</b>																									
16. SUPPLEMENTARY NOTATION																									
17. COSATI CODES <table border="1"><tr><td>FIELD</td><td>GROUP</td><td>SUB. GR.</td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td></tr></table>			FIELD	GROUP	SUB. GR.										18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)										
FIELD	GROUP	SUB. GR.																							
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			<table border="1"><tr><td colspan="2">Approved for</td></tr><tr><td>DTIC GRA&amp;I</td><td><input checked="" type="checkbox"/></td></tr><tr><td>DTIC TAB</td><td><input checked="" type="checkbox"/></td></tr><tr><td>Unannounced</td><td><input type="checkbox"/></td></tr><tr><td>Justification</td><td></td></tr><tr><td colspan="2">By _____</td></tr><tr><td colspan="2">Distribution/</td></tr><tr><td colspan="2">Availability Codes</td></tr><tr><td>Avail and/or</td><td></td></tr><tr><td>Dist</td><td>Special</td></tr></table>			Approved for		DTIC GRA&I	<input checked="" type="checkbox"/>	DTIC TAB	<input checked="" type="checkbox"/>	Unannounced	<input type="checkbox"/>	Justification		By _____		Distribution/		Availability Codes		Avail and/or		Dist	Special
Approved for																									
DTIC GRA&I	<input checked="" type="checkbox"/>																								
DTIC TAB	<input checked="" type="checkbox"/>																								
Unannounced	<input type="checkbox"/>																								
Justification																									
By _____																									
Distribution/																									
Availability Codes																									
Avail and/or																									
Dist	Special																								
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <b>UNCLASSIFIED/UNLIMITED</b> <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>			21. ABSTRACT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>																						
22a. NAME OF RESPONSIBLE INDIVIDUAL <b>MAJOR AMOS L. OTIS</b>			22b. TELEPHONE NUMBER (Include Area Code) <b>(202) 767-4971</b>		22c. OFFICE SYMBOL <b>XOT</b>																				

17.

AD A 154 337

## PREFACE

The United States Air Force Summer Faculty Research Program (USAF-SFRP) is a program designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis for a ten-week assignment during the summer intersession period to perform research at Air Force laboratories/centers. Each assignment is in a subject area and at an Air Force facility mutually agreed upon by the faculty members and the Air Force. In addition to compensation, travel and cost of living allowances are also paid. The USAF-SFRP is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, United States Air Force, and is conducted by the Southeastern Center for Electrical Engineering.

The specific objectives of the 1984 USAF-SFRP are:

- (1) To provide a productive means for Scientists and Engineers holding Ph.D. degrees to participate in research at the Air Force Weapons Laboratory;
- (2) To stimulate continuing professional association among the Scholars and their professional peers in the Air Force;
- (3) To further the research objectives of the United States Air Force;
- (4) To enhance the research productivity and capabilities of Scientists and Engineers especially as these relate to Air Force technical interests.

During the summer of 1984, 152 faculty members participated. These researchers were assigned to 25 USAF laboratories/centers across the country. This three volume document is a compilation of the final reports written by the assigned faculty members about their summer research efforts.

UNITED STATES AIR FORCE  
SUMMER FACULTY RESEARCH PROGRAM

1984

PROGRAM MANAGEMENT REPORT  
SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

Volume II of III

Program Directors, SCEEE  
Warren D. Peele  
Earl L. Steele

Program Manager, AFOSR  
Major Amos L. Otis

Submitted to  
Air Force Office of Scientific Research  
Bolling Air Force Base  
Washington, DC

December 1984

## PREFACE

The United States Air Force Summer Faculty Research Program (USAF-SFRP) is a program designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis for a ten-week assignment during the summer intersession period to perform research at Air Force laboratories/centers. Each assignment is in a subject area and at an Air Force facility mutually agreed upon by the faculty members and the Air Force. In addition to compensation, travel and cost of living allowances are also paid. The USAF-SFRP is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, United States Air Force, and is conducted by the Southeastern Center for Electrical Engineering.

The specific objectives of the 1984 USAF-SFRP are:

- (1) To provide a productive means for Scientists and Engineers holding Ph.D. degrees to participate in research at the Air Force Weapons Laboratory;
- (2) To stimulate continuing professional association among the Scholars and their professional peers in the Air Force;
- (3) To further the research objectives of the United States Air Force;
- (4) To enhance the research productivity and capabilities of Scientists and Engineers especially as these relate to Air Force technical interests.

During the summer of 1984, 152 faculty members participated. These researchers were assigned to 25 USAF laboratories/centers across the country. This three volume document is a compilation of the final reports written by the assigned faculty members about their summer research efforts.

# LIST OF PARTICIPANTS

## NAME/ADDRESS

## DEGREE, SPECIALTY, LABORATORY ASSIGNED

Dr. Annalingam Anandarajah  
Assistant Professor  
S.D. School of Mines & Tech.  
Civil Engineering Dept.  
Rapid City, SD 57701  
(605) 394-2443

Degree: Ph.D., Civil Engineering  
1982  
Specialty: Geotechnical Engineering  
Assigned: ESC

Dr. Gloria Anderson  
Chairman  
Morris Brown College  
Chemistry Dept.  
Atlanta, GA 30314  
(404) 525-7831

Degree: Ph.D., Chemistry, 1968  
Specialty: Organic Chemistry  
Assigned: RPL

Dr. Richard Anderson  
Professor  
University of Missouri  
Physics Department  
Rolla, MO 65401  
(314) 341-4341

Degree: Ph.D., Physics, 1959  
Specialty: Optics, Atomic and Molecular  
Physics  
Assigned: APL

Dr. Deborah Armstrong  
Assistant Professor  
University of Texas  
Life Sciences  
San Antonio, TX 78285  
(512) 691-4458

Degree: Ph.D., Biopsychology, 1982  
Specialty: Neurophysiology  
Assigned: SAM

Dr. Francesco Bacchialoni  
Associate Professor  
University of Lowell  
Dept. of Electrical Eng.  
Lowell, MA 01854  
(617) 452-5000

Degree: Ph. D., Engineering, 1946  
Specialty: Control Systems, Digital  
Signal Processing,  
Microprocessors  
Assigned: GL

Dr. John Bahng  
Associate Professor  
Northwestern University  
Dept. of Physics & Astronomy  
Evanston, IL 60201  
(312) 492-7527

Degree: Ph.D., Astronomy, 1957  
Specialty: Astronomical Instrumentation,  
Computer Programming  
Assigned: GL



List of Participants (continued: page 2)

Dr. James Baird  
Professor  
Brown University  
Chemistry & Physics Dept.  
Providence, RI 02912  
(401) 863-3325

Degree: Ph.D., Physics, 1959  
Specialty: Spectroscopy-laser,  
Microwave  
Assigned: GL

Dr. Mukul Banerjee  
Professor  
Meharry Medical College  
Dept. of Physiology  
Nashville, TN 37208  
(615) 327-6288

Degree: Ph.D., Animal Physiology,  
1964  
Specialty: Respiratory Physiology,  
Environmental Physiology  
Assigned: SAM

Dr. Alan Bentley  
Assistant Professor  
Eastern Montana College  
Dept. of Physics  
Billings, MT 59101  
(406) 657-2026

Degree: Ph.D., Infrared  
Astrophysics, 1980  
Specialty: Infrared Physics and  
Astrophysics  
Assigned: GL

Dr. Richard Bernhard  
Professor  
North Carolina State Univ.  
Industrial Engr. Dept.  
Raleigh, NC 27695  
(919) 737-2362

Degree: Ph.D., Operations Research,  
1961  
Specialty: Engineering and Managerial  
Economics; Decision Anal.  
Assigned: BRMC

Dr. Albert Biggs  
Professor  
University of Kansas  
Electrical Engr. Dept.  
Lawrence, KS 66045  
(913) 864-4836

Degree: Ph.D., Electrical Engr.,  
1968  
Specialty: Microwaves and Antennas  
Assigned: AL

Dr. Louis Buckalew  
Associate Professor  
Alabama A&M University  
Dept. of Psychology  
Normal, AL 35762  
(205) 859-7451

Degree: Ph.D., Phys. Psychology,  
1984  
Specialty: Physiological Psychology  
Assigned: LMDC

Mr. Mike Burlakoff  
Assistant Professor  
Southwest Mo. State Univ.  
Computer Science Dept.  
Springfield, MO 65804  
(417) 836-5930

Degree: M.S., Math/Computer Science,  
1965  
Specialty: Languages and Environments  
Assigned: AL

List of Participants (continued: page 3)

Dr. Myron Calhoun  
Associate Professor  
Kansas State University  
Computer Science Dept.  
Manhattan, KS 66506  
(913) 532-6350

Degree: Ph. D., Electrical Engr.,  
1967  
Specialty: Digital Hardware and  
Software  
Assigned: AD

Dr. Jeya Chandra  
Assistant Professor  
Pennsylvania State Univ.  
Dept. of Ind. & Mgmt. Engr.  
University Park, PA 16802  
(814) 863-2358

Degree: Ph. D., Ind. Eng. and  
Operations Rsch., 1980  
Specialty: Stochastic Processes  
Assigned: SAM

Dr. Do Chang  
Associate Professor  
Averett College  
Chemistry Department  
Danville, VA 24541  
(804) 793-7811

Degree: Ph. D., Chemistry, 1968  
Specialty: Chemical Kinetics, Phase  
Transitions  
Assigned: AD

Dr. Huei-huang Chiu  
Professor  
University of Illinois  
Dept. of Mechanical Engr.  
Chicago, IL 60680  
(312) 996-3426

Degree: Ph. D., Aeronautical Eng.,  
1962  
Specialty: Combustion, Fluid  
Mechanics, and Propulsion  
Assigned: APL

Dr. Philip Chong  
Assistant Professor  
Dept. of Industrial Engr.  
North Dakota State Univ.  
Fargo, ND 58105  
(701) 237-7223

Degree: Ph. D., Ind. Eng. and  
Operations Rsch., 1977  
Specialty: Computerized Resource  
Planning and Scheduling  
Assigned: LMC

Dr. Louis Chow  
Assistant Professor  
Washington State University  
Mechanical Engr. Dept.  
Pullman, WA 99164  
(509) 335-1327

Degree: Ph. D., Mechanical Engr.,  
1978  
Specialty: Heat Transfer, Fluid  
Dynamics  
Assigned: APL

Dr. David Chung  
Professor  
Howard University  
Department of Physics  
Washington, DC 20059  
(202) 636-7903

Degree: Ph. D., Solid State  
Physics, 1966  
Specialty: Fiber Optics Sensors Ultra-  
sound, Solid State Elect.  
Assigned: FJSRL

List of Participants (continued: page 4)

Dr. David Cohoon  
Professor  
Temple University  
Dept. of Mathematics  
Philadelphia, PA 19122  
(215) 787-7284

Degree: Ph. D., Mathematics, 1969  
Specialty: Partial Differential  
Equations  
Assigned: SAM

Dr. Frank Colby, Jr.  
Assistant Professor  
University of Lowell  
Dept. of Earth Sciences  
Lowell, MA 01854  
(617) 452-2551

Degree: Ph. D., Meteorology, 1983  
Specialty: Boundary-layer  
Meteorology  
Assigned: GL

Dr. Robert Colclaser, Jr.  
Professor & Chairman  
University of Pittsburgh  
Electrical Engineer  
Pittsburgh, PA 15261  
(412) 624-5391

Degree: Ph. D., Sci., Elec. Engr.,  
1968  
Specialty: High Power Arcs, Circuit  
Breaker Design and Test,  
Electrical Transients  
Assigned: WL

Dr. Gregory Corso  
Assistant Professor  
Georgia Inst. of Tech.  
School of Psychology  
Atlanta, GA 30332  
(404) 894-2680

Degree: Ph. D., Engr. Psychology,  
1978  
Specialty: Human Factors, Human  
Performance  
Assigned: AEDC

Dr. Robert Courter  
Associate Professor  
Louisiana State University  
Mechanical Engr. Dept.  
Baton Rouge, LA 70803  
(504) 388-5792

Degree: Ph. D., Aerospace Engr.,  
1965  
Specialty: Aerodynamics, Gasdynamics,  
Blast Waves  
Assigned: AD

Dr. John Cyranski  
Assistant Professor  
Clark College  
Physics Department  
Atlanta, GA 30314  
(404) 681-3080

Degree: Ph. D., Physics, 1974  
Specialty: Math. Physics and  
Information Theory  
Assigned: GL

Dr. Subramaniam Deivanayagam  
Associate Professor  
University of Texas  
Dept. of Industrial Engr.  
Arlington, TX 76109  
(817) 273-3092

Degree: Ph. D., Industrial Engr.,  
1973  
Specialty: Ergonomics/Human Factors  
Assigned: HRL/Wright-Patterson

List of Participants (continued: page 5)

Dr. Hermann Donnert  
Professor  
Kansas State University  
Dept. of Nuclear Engr.  
Manhattan, KS 66506  
(913) 532-5960

Degree: Ph. D., Mathematics and  
Physics, 1951  
Specialty: Plasma Physics, Radiation  
Physics  
Assigned: FJSRL

Dr. Robert Dorman  
Assistant Professor  
Kent State University  
Dept. of Biological Sci.  
Kent, OH 44242  
(216) 672-3613

Degree: Ph. D., Physiological  
Chemistry, 1976  
Specialty: Neurochemistry  
Assigned: SAM

Dr. George Doyle, Jr.  
Associate Professor  
University of Dayton  
Mechanical Engr. Dept.  
Dayton, OH 45469  
(513) 229-2835

Degree: Ph. D., Mechanical Engr.,  
1973  
Specialty: Dynamics  
Assigned: FDL

Dr. Eric Drumm  
Assistant Professor  
University of Tennessee  
Dept. of Civil Engr.  
Knoxville, TN 37996  
(615) 974-7715

Degree: Ph. D., Civil Engineering,  
1983  
Specialty: Geotechnical Engineering  
Assigned: ESC

Dr. Charles Drummond, III  
Associate Professor  
Ohio State University  
Dept. of Ceramic Engr.  
Columbus, OH 43210  
(614) 422-2960

Degree: Ph. D., Applied Physics,  
1974  
Specialty: Glass Structure and  
Properties  
Assigned: ML

Dr. Delcie Durham  
Assistant Professor  
University of Vermont  
Civil & Mech. Engr. Dept.  
Burlington, VT 05405  
(802) 656-3320

Degree: Ph. D., Mechanical Engr.,  
1981  
Specialty: Metal Behavior at High  
Strains and Strain Rates  
Assigned: ML

Dr. Terrence Dwan  
Associate Professor  
The Citadel  
Dept. of Elect. Engr.  
Charleston, SC 29409  
(803) 792-5057

Degree: Ph. D., Electrical Engr.  
1975  
Specialty: Large Scale Systems,  
Modelling, Controls  
Assigned: AD

List of Participants (continued: page 6)

Dr. Franklin Eastep  
Professor  
University of Dayton  
Dept. of Aerospace Engr.  
Dayton, OH 45469  
(513) 229-2241

Degree: Ph. D., Aeronautics, 1968  
Specialty: Aeroelasticity  
Assigned: FDL

Dr. James Eberhart  
Professor  
University of Colorado  
Chemistry Department  
Colorado Springs, CO 80933  
(303) 593-3284

Degree: Ph. D., Chemistry, 1963  
Specialty: Physical and Surface  
Chemistry  
Assigned: FJSRL

Dr. Frederick Eisler  
Associate Professor  
College of Staten Island  
Dept. of Applied Science  
Staten Island, NY 10301  
(212) 390-7973

Degree: Ph. D., Physics, 1970  
Specialty: Particle Physics,  
Holography, Accelerators  
Assigned: WL

Dr. Emory Ensore, Jr.  
Associate Professor  
Penn. State University  
Dept. of Ind. & Mgmt. Syst. Eng.  
University Park, PA 16802  
(814) 863-2353

Degree: Ph. D., Ind. Engr., 1972  
Specialty: Industrial Engineering  
Assigned: HRL/Wright-Patterson

Dr. John Erdei  
Assistant Professor  
University of Dayton  
Dept. of Physics  
Dayton, OH 45469  
(513) 229-2727

Degree: Ph. D., Physics, 1983  
Specialty: Many Body Theory, Critical  
Phenomena  
Assigned: APL

Dr. Adly Fam  
Associate Professor  
State University of New York  
Dept. of Electrical & Comp. Eng.  
Buffalo, NY 14260  
(716) 636-2422

Degree: Ph. D., Electrical Engr.,  
1977  
Specialty: Digital Signal Processing  
and System Theory  
Assigned: RADC

Dr. Bruce R. Feiring  
Assistant Professor  
University of Minnesota  
Dept. of Mgmt. Sciences  
Minneapolis, MN 55455  
(612) 376-1376

Degree: Ph. D., Industrial Engr.,  
1979  
Specialty: Optimization  
Assigned: HRL/Brooks

List of Participants (continued: page 7)

Dr. William Feld  
Associate Professor  
Wright State University  
Dept. of Chemistry  
Dayton, OH 45435  
(513) 873-2511

Degree: Ph. D., Chemistry, 1971  
Specialty: Organic Chemistry  
Assigned: ML

Dr. Dennis Flentge  
Assistant Professor  
Cedarville College  
Dept. of Math/Sci.  
Cedarville, OH 45314  
(513) 766-2211

Degree: Ph. D., Physical Chemistry,  
1974  
Specialty: Catalysis, IR Spectroscopy  
Assigned: APL

Dr. Cynthia Ford  
Assistant Professor  
Jackson State University  
Psychology Department  
Jackson, MS 39217  
(601) 968-2371

Degree: Ph. D., Educ. Psychology,  
1979  
Specialty: General Psychology,  
Statistics  
Assigned: HRL/Brooks

Dr. Eddie Fowler  
Associate Professor  
Kansas State University  
Electrical Engr. Dept.  
Manhattan, KS 66056  
(913) 532-5600

Degree: Ph. D., Elec. Eng., 1969  
Specialty: C<sup>3</sup>I Modeling and Sim-  
ulation Nuclear Degrad-  
ation of Comm. Networks  
Assigned: WL

Dr. Basil Gala  
Professor  
California State University  
Division of Engineering  
Chico, CA 95929  
(916) 895-5374

Degree: Ph. D., Electrical Engr.,  
1973  
Specialty: Computers-Pattern  
Recognition  
Assigned: RADC

Dr. Barry Ganapol  
Associate Professor  
University of Arizona  
Dept. of Nuclear & Energy Engr.  
Tucson, AZ 85721  
(602) 621-4728

Degree: Ph. D., Nuclear Science,  
1971  
Specialty: Particle Transport Theory  
Assigned: RADC

Dr. David Gilliam  
Assistant Professor  
Texas Tech. University  
Dept. of Mathematics  
Lubbock, TX 79409  
(806) 742-2566

Degree: Ph. D., Mathematics, 1977  
Specialty: Applied Partial  
Differential Equations  
Assigned: RADC

List of Participants (continued: page 8)

Dr. Larry Glasgow  
Assistant Professor  
Kansas State University  
Chemical Engr. Dept.  
Manhattan, KS 66506  
(913) 532-5585

Degree: Ph. D., Chemical Engr.,  
1977  
Specialty: Fluid Mechanics, Transport  
Phenomena, Laser-Doppler  
Velocity  
Assigned: FJSRL

Dr. Sallie Gordon  
Assistant Professor  
Wright State University  
Dept. of Psychology  
Dayton, OH 45435  
(513) 873-2391

Degree: Ph. D., Psychology, 1982  
Specialty: Social/Cognitive  
Psychology  
Assigned: HRL/Wright-Patterson

Dr. Thomas Graham  
Professor  
University of Dayton  
Physics Dept.  
Dayton, OH 45469  
(513) 229-2329

Degree: Ph. D., Physics, 1967  
Specialty: Solid State/Surface  
Physics/ Magnetic  
Resonance  
Assigned: ML

Dr. Edward Greco, Jr.  
Assistant Professor  
University of Miami  
Biomedical Engr. Dept.  
Coral Gables, FL 33124  
(305) 284-2442

Degree: Ph. D., Elec./Bioengr.,  
1976  
Specialty: Digital Signal Processing,  
Biosystem Analysis and  
Ventilatory Control  
Assigned: SAM

Dr. Ronald Greene  
Associate Professor  
University of New Orleans  
Physics Department  
New Orleans, LA 70148  
(504) 286-6714

Degree: Ph. D., Physics, 1974  
Specialty: Plasma Spectroscopy  
Assigned: AL

Dr. Paul Griesacker  
Associate Professor  
Gannon University  
Department of Physics  
Erie, PA 16541  
(814) 871-7649

Degree: Ph. D., Physics, 1963  
Specialty: Physical Optics, Coherent  
Radiation  
Assigned: AL

Dr. Thomas Gullledge, Jr.  
Assistant Professor  
Louisiana State University  
Dept. of Quantitative Bus. Anal.  
Baton Rouge, LA 70803  
(504) 388-2126

Degree: Ph. D., Eng. Mgmt., 1981  
Specialty: Management Science  
Assigned: BRMC

List of Participants (continued: page 9)

Dr. Vijay Gupta  
Associate Professor  
Central State University  
Chemistry Department  
Wilberforce, OH 45384  
(513) 376-6423

Degree: Ph. D., Chemistry, 1969  
Specialty: Physical Chemistry  
Assigned: ML

Dr. Hendrik Hamaka  
Professor  
University of Pennsylvania  
Chemistry Dept.  
Philadelphia, PA 19104  
(215) 898-8303

Degree: Ph. D., Theoretical Chem.,  
1956  
Specialty: Quantum Chemistry,  
Theoretical Chemistry  
Assigned: FJSRL

Dr. Arthur Harriman  
Professor  
Oklahoma State University  
Dept. of Psychology  
Stillwater, OK 74078  
(405) 624-6561

Degree: Ph. D., Exp. Psychology,  
1952  
Specialty: Physiological Psychology  
Assigned: HRL/Williams

Dr. Doyle Hasty  
Assistant Professor  
Motlow State Community College  
Dept. of Engineering  
Tullahoma, TN 37388  
(615) 455-8511

Degree: M.S., Engineering Admn.,  
1974  
Specialty: Electronics, Physics,  
Instrumentation, Computers,  
High-altitude Engine Test.  
Assigned: AEDC

Dr. Albert Havener  
Assistant Professor  
University of Dayton  
Dept. of Mechanical Engr.  
Dayton, OH 45469  
(513) 229-2835

Degree: Ph. D., Aerospace Engr.,  
1983  
Specialty: Applied Aero Optical  
Measuring Techniques  
Assigned: FDL

Dr. Peter Hierl  
Professor  
Kansas University  
Chemistry Department  
Lawrence, KS 66045  
(913) 864-3019

Degree: Ph. D., Physical Chemistry  
1964  
Specialty: Gas-phase Kinetics  
Assigned: GL

Dr. Paul Hoffman  
Assistant Professor  
Villanova University  
Dept. of Civil Engineering  
Villanova, PA 19085  
(215) 645-4960

Degree: Ph. D., Civil Engineering  
1982  
Specialty: Structural Eng. and Solid  
Mechanics  
Assigned: ESC



List of Participants (continued: page 10)

Dr. Brian Holmes Assistant Professor San Jose State University Dept. of Physics San Jose, CA 95912 (408) 277-2361	<u>Degree:</u> Ph. D., Physics, 1980 <u>Specialty:</u> Solid State/Low Temperature; Magnetic Resonance <u>Assigned:</u> RADC
Dr. Gwendolyn Howze Associate Professor Texas Southern University Dept. of Biology Houston, TX 77004 (713) 527-7005	<u>Degree:</u> Ph. D., Molecular Biology, 1974 <u>Specialty:</u> Cell Biology/Chromatin, Electron Microscopy, Tissue Cult <u>Assigned:</u> AMRL
Dr. Chen-Chi Hsu Professor University of Florida Dept. of Engineering Sci. Gainesville, FL 32611 (904) 392-0961	<u>Degree:</u> Ph. D., Eng. Mech., 1965 <u>Specialty:</u> Applied Mechanics, Computational Aerodynamics <u>Assigned:</u> AD
Dr. Mario Innocenti Assistant Professor Auburn University Aerospace Engr. Dept. Auburn, AL 36849 (205) 926-4874	<u>Degree:</u> Ph. D., Aeronautics Astronautics, 1983 <u>Specialty:</u> Flight Dynamics Handling Qualities Optimal Control <u>Assigned:</u> FDL
Dr. Kakkattukuzhy Isaac Assistant Professor University of Missouri Dept. of Mech. & Aero. Engr. Rolla, MO 65401 (314) 341-4626	<u>Degree:</u> Ph. D., Aerospace Eng., 1982 <u>Specialty:</u> Fluid Mechanics <u>Assigned:</u> APL
Dr. Robert Jackson, Jr. Assistant Professor University of Massachusetts Dept. of Elec. & Comp. Engr. Amherst, MA 01003 (413) 545-1386	<u>Degree:</u> Ph. D., Elec. Eng., 1982 <u>Specialty:</u> Electromagnetics, Active Devices for Microwave and Millimeter Wave Integrated Circuits <u>Assigned:</u> RADC
Dr. Vinod Jain Assistant Professor University of Dayton Dept. of Mechanical Engr. Dayton, OH 45469 (513) 229-2835	<u>Degree:</u> Ph. D., Mech. Eng., 1980 <u>Specialty:</u> Materials and Manufactur- ing, Design, and Tri- biology <u>Assigned:</u> ML

List of Participants (continued: page 11)

Dr. Bruce Janson  
Assistant Professor  
Carnegie Mellon University  
Dept. of Civil Engr.  
Pittsburgh, PA 15213  
(412) 578-3866

Degree: Ph. D., Civil Eng., 1981  
Specialty: Engineering Planning and  
Management  
Assigned: LMC

Dr. Charles Jones  
Associate Professor  
North Carolina Central  
Dept. of Physics  
Durham, NC 27707  
(919) 684-6452

Degree: Ph. D., Physics, 1977  
Specialty: Lasers and their  
Applications, Optics,  
Electro-Optics  
Assigned: AD

Dr. Walter Jones  
Assistant Professor  
University of Tennessee  
Dept. of Engr. Sci. Mech.  
Knoxville, TN 37996  
(615) 974-7684

Degree: Ph. D., Engr. Mechanics,  
1982  
Specialty: Mechanical Behavior of  
Composite Materials  
Assigned: FDL

Dr. Robert Kallman  
Professor  
North Texas State University  
Dept. of Mathematics  
Denton, TX 76203  
(817) 565-2155

Degree: Ph. D., Mathematics, 1968  
Specialty: Analysis  
Assigned: AD

Dr. William Kane, Jr.  
Associate Professor  
Western Carolina University  
Management Dept.  
Cullowhee, NC 28723  
(704) 227-7401

Degree: Ph. D., Organ. Behavior,  
1977  
Specialty: Organizational Behavior-  
Strategic Management  
Assigned: HRL/Wright-Patterson

Dr. Yong Kim  
Instructor  
McNeese State University  
Civil & Mechanical Engr. Dept.  
Lake Charles, LA 70609  
(318) 477-2520

Degree: Ph. D., Civil Engineering  
1984  
Specialty: Geotechnical Engineering  
Assigned: ESC

Dr. Ronald Kline  
Assistant Professor  
University of Oklahoma  
School of Aero. Mech., & Nuc. Eng.  
Norman, OK 73019  
(405) 325-5011

Degree: Ph. D., Mech, and Material  
Sci., 1978  
Specialty: Nondestructive Testing,  
Mechanics of Composite  
Materials  
Assigned: ML

List of Participants (continued: page 12)

Dr. Kent Knaebel  
Assistant Professor  
Ohio State University  
Chemical Engineering Dept.  
Columbus, OH 43210  
(614) 422-2508

Degree: Ph. D., Chemical Engr.,  
1980  
Specialty: Adsorption, Ion-exchange,  
Separations  
Assigned: SAM

Dr. David Kohfeld  
Professor  
Southern Illinois University  
Dept. of Psychology  
Edwardsville, IL 62006  
(618) 692-2582

Degree: Ph. D., Exper. Psychology,  
1966  
Specialty: Methods of Psychological  
Inquiry  
Assigned: HRL/Williams

Dr. Gabriel Kojoian  
Associate Professor  
University of Wisconsin  
Dept. of Physics/Astronomy  
Eau Claire, WI 54701  
(715) 836-3148

Degree: Ph. D., Physics, 1966  
Specialty: Radio Astronomy  
Assigned: GL

Dr. Arthur Kovitz  
Professor  
Northwestern University  
Mechanical Engr. Dept.  
Evanston, IL 60201  
(312) 492-7066

Degree: Ph. D., Aero Eng., 1957  
Specialty: Fluid Mechanics  
Assigned: WL

Dr. Madakasira Krishna  
Associate Professor  
South Carolina State College  
Dept. of Math. & Comp. Sci.  
Orangeburg, SC 29117  
(803) 536-7121

Degree: Ph. D., Applied Math.,  
1974  
Specialty: Computational Fluid  
Dynamics  
Assigned: AEDC

Dr. Raj Krishnan  
Associate Professor  
North Texas State University  
Physics Dept.  
Denton, TX 76203  
(817) 565-3284

Degree: Ph. D., Physics, 1966  
Specialty: Nuclear Physics, Solid  
State Physics  
Assigned: SAM

Dr. William Kyros  
Associate Professor  
University of Lowell  
Mechanical Eng. Dept.  
Lowell, MA 01854  
(617) 452-5000

Degree: Ph. D., Education, 1980  
Specialty: Mechanical Behavior of  
Materials  
Assigned: ML

List of Participants (continued: page 13)

Dr. David Lai  
Professor  
University of Vermont  
Electrical Engr. & Comp. Sci.  
Burlington, VT 05405  
(802) 656-3330

Degree: Ph. D., Elec. Eng., 1960  
Specialty: Signal Processing, Radar  
Signal Processing, Image  
Processing  
Assigned: RADC

Dr. Charles Lardent, Jr.  
Associate Professor  
Auburn University  
Dept. of Management  
Montgomery, AL 36117  
(205) 271-9478

Degree: Ph. D., Organ. Behavior,  
1979  
Specialty: Leadership, Motivation,  
and Psychometric  
Assessment  
Assigned: LMDC

Dr. Nabil Lawandy  
Assistant Professor  
Brown University  
Engineering Dept.  
Providence, RI 02912  
(401) 863-2755

Degree: Ph. D., Chem. Physics,  
1980  
Specialty: Lasers, Spectroscopy,  
Lasers Atom Interactions  
Assigned: GL

Dr. E. Miller Layton, Jr.  
Assoc. Professor  
Sterling College  
Dept. of Chemistry & Applied Math.  
Sterling, KS 67579  
(316) 278-2173

Degree: Ph. D., Chem. Physics,  
1962  
Specialty: Molecular Spectroscopy,  
Theoretical Molecular  
Calculations  
Assigned: RPL

Dr. Evelyn Leggette  
Associate Professor  
Jackson State University  
Dept. of Elem./Early Child/Reading  
Jackson, MS 39217  
(601) 968-2186

Degree: Ph. D., Curriculum and  
Inst., 1976  
Specialty: Reading Ed., Dev. Ed.,  
Prog. Dev. and  
Implementation Curriculum  
Assigned: LMC

Dr. Meng Liou  
Associate Professor  
University of Michigan  
Aerospace Engr. Dept.  
Ann Arbor, MI 48104  
(313) 764-3310

Degree: Ph. D., Aerospace Eng.,  
1977  
Specialty: Computational Fluid  
Dynamics, Gas Dynamics,  
Unsteady Flow  
Assigned: FDL

Dr. David Lohman  
Assistant Professor  
University of Iowa  
College of Education  
Iowa City, IA  
(319) 353-5961

Degree: Ph. D., Educ. Psychology,  
1979  
Specialty: Information Processing  
Theories of Aptitude for  
Learning  
Assigned: HRL/Brooks

List of Participants (continued: page 14)

Dr. Lonnie Ludeman  
Professor  
New Mexico State University  
Dept. of Elect. & Comp. Engr.  
Las Cruces, NM 88003  
(505) 646-1321

Degree: Ph. D., Elec. Eng., 1968  
Specialty: Digital Signal Processing,  
Statistical Communication  
Theory  
Assigned: RADC

Dr. Larry Ludwick  
Professor  
Tuskegee Institute  
Dept. of Chemistry  
Tuskegee Institute, AL 36088  
(205) 727-8836

Degree: Ph. D., Inorganic Chem.,  
1969  
Specialty: Transition Metal Chem.,  
Siloxane Polymers  
Assigned: ML

Dr. Robert MacCallum  
Associate Professor  
Ohio State University  
Psychology Dept.  
Columbus, OH 43210  
(614) 422-1030

Degree: Ph. D., Psychology, 1974  
Specialty: Quantitative Psychology  
Assigned: AMRL

Dr. William McCormick  
Associate Professor  
Wright State University  
Dept. of Electrical Engr.  
Dayton, OH 45435  
(513) 873-2849

Degree: Ph. D., Electrical Engr.  
1967  
Specialty: Radar, Communications and  
Electromagnetics  
Assigned: AL

Dr. Odis McDuff  
Professor  
University of Alabama  
Dept. of Electrical Engr.  
University, AL 35486  
(205) 348-6351

Degree: Ph. D., Electrical Engr.,  
1966  
Specialty: Lasers and Optics,  
Electromagnetics  
Assigned: SAM

Dr. Bernard McIntyre  
Associate Professor  
University of Houston  
Electronics Dept.  
Houston, TX 77004  
(713) 749-4753

Degree: Ph. D., Solid State  
Physics, 1970  
Specialty: Diamagnetism in Metals  
Assigned: GL

Dr. Richard Miers  
Associate Professor  
Indiana University  
Department of Physics  
Fort Wayne, IN 46805  
(219) 482-5693

Degree: Ph. D., Atomic Physics,  
1970  
Specialty: Atom Physics  
Assigned: APL

List of Participants (continued: page 15)

Dr. John Minor  
Assistant Professor  
University of Oklahoma  
Elect. Engr. & Comp. Sci. Dept.  
Norman, OK 73069  
(405) 325-4721

Degree: Ph. D., Computer Science,  
1979  
Specialty: Artificial Intelligence  
Assigned: RADC

Dr. Charles Mitchell  
Professor  
Colorado State University  
Dept. of Mechanical Engr.  
Fort Collins, CO 80523  
(303) 491-6558

Degree: Ph. D., Aerospace and  
Mechanical Science, 1967  
Specialty: Combustion Instability,  
Combustion, Gas Dynamics  
Assigned: RPL

Dr. Don Mittleman  
Professor  
Oberlin College  
Dept. of Mathematics  
Oberlin, OH 44074  
(216) 775-8385

Degree: Ph. D., Mathematics, 1951  
Specialty: Geometry, Analysis, Mech.  
Assigned: FDL

Dr. Dale Moses  
Associate Professor  
San Diego State University  
Aerospace & Mechan. Engr. Dept.  
San Diego, CA 92182  
(619) 265-5764

Degree: Ph. D., Aerospace Engr.,  
1981  
Specialty: Wind Tunnel Testing,  
Aerodynamics  
Assigned: FDL

Dr. Kevin Mossholder  
Associate Professor  
Auburn University  
Management Department  
Auburn, AL 36849  
(205) 826-4071

Degree: Ph. D., Ind./Organ.  
Psychology, 1978  
Specialty: I/O Topics: Organizational  
Behavior, Human Resources  
Mgmt.  
Assigned: LMDC

Dr. James Mrotek  
Associate Professor  
Meharry Medical College  
Dept. of Physiology  
Nashville, TN 37208  
(615) 327-6979

Degree: Ph. D., Biology, 1973  
Specialty: Endocrine Cell  
Intracellular Exchanges  
Assigned: SAM

Dr. Richard Murphy  
Professor  
University of Missouri  
Department of Physics  
Kansas City, MO 64110  
(816) 276-1604

Degree: Ph. D., Physics, 1968  
Specialty: Statistical Mechanics;  
Theory of Liquids and  
Solids  
Assigned: FJSRL

List of Participants (continued: page 16)

Dr. Lena Myers Professor Jackson State University Department of Sociology Jackson, MS 39217 (601) 968-2591	<u>Degree:</u> Ph. D., Sociology, 1973 <u>Specialty:</u> Social Psychology <u>Assigned:</u> SAM
Dr. Datta Naik Chairman & Assoc. Professor Monmouth College Dept. of Chemistry West Long Branch, NJ 07764 (201) 222-6600	<u>Degree:</u> Ph. D., Chemistry, 1972 <u>Specialty:</u> Analytical/Inorganic Chemistry <u>Assigned:</u> ESC
Dr. Stephan Nix Assistant Professor Syracuse University Dept. of Civil Engr. Syracuse, NY 13210 (315) 423-2311	<u>Degree:</u> Ph. D., Environmental Eng., 1982 <u>Specialty:</u> Environmental Resources Mgmt., Operations Rsch., Economics, Hydrology <u>Assigned:</u> OEHL
Dr. William Norton Associate Professor Southeastern La. University Biology Department Hammond, LA 70402 (504) 549-2173	<u>Degree:</u> Ph. D., Cell Biology, 1975 <u>Specialty:</u> Cytopathology, Cellular Ultrastructure <u>Assigned:</u> AMRL
Dr. Kendall Nygard Associate Professor North Dakota State University Dept. of Mathematical Sci. Fargo, ND 58103 (701) 237-8178	<u>Degree:</u> Ph. D., Oper. Rsch., 1978 <u>Specialty:</u> Optimization, Simulation Modeling, Computer Graphics <u>Assigned:</u> LC
Dr. Robert O'Connell Assistant Professor University of Missouri Dept. of Electrical & Comp. Eng. Columbia, MO 65211 (314) 882-8373	<u>Degree:</u> Ph. D., Elec. Eng., 1975 <u>Specialty:</u> Physical and Quantum Electronics, and Applied Optics <u>Assigned:</u> FJSRL
Dr. William Pardo Associate Professor University of Miami Physics Department Coral Gables, FL 33124 (305) 284-2323	<u>Degree:</u> Ph. D., Physics, 1957 <u>Specialty:</u> Experimental Plasma Physics <u>Assigned:</u> AD

List of Participants (continued: page 17)

Dr. Martin Patt  
Associate Professor  
University of Lowell  
Department of Elect. Engr.  
Lowell, MS 01854  
(617) 452-5000

Degree: M.S., Electrical Engr.,  
1964  
Specialty: Computer Science  
Assigned: GL

Dr. James Patterson  
Professor  
S.D. School of Mines/Tech.  
Physics Dept.  
Rapid City, SD 57701  
(605) 394-2361

Degree: Ph. D., Physics, 1962  
Specialty: Theoretical Solid State  
Physics  
Assigned: AL

Dr. M. Carr Payne, Jr.  
Professor  
Georgia Inst. of Tech.  
Dept. of Psychology  
Atlanta, GA 30332  
(404) 894-2681

Degree: Ph. D., Psychology, 1951  
Specialty: Experimental Psychology,  
Psychoacoustics, Eng. Psy.  
Assigned: AEDC

Dr. William Perrizo  
Associate Professor  
North Dakota State University  
Dept. of Computer Science  
Fargo, ND 58105  
(701) 237-7248

Degree: Ph. D., Dynamical Systems,  
1972  
Specialty: Distributed Database  
Systems  
Assigned: ESD

Dr. Boake Plessy  
Professor  
Dillard University  
Chemistry Department  
New Orleans, LA 70122  
(504) 283-8822

Degree: Ph. D., Physical Chem.,  
1974  
Specialty: Polymers and Biopolymers,  
Analytical Chemistry  
Assigned: SAM

Dr. Kuldip Rattan  
Associate Professor  
Wright State University  
Dept. of Systems Eng.  
Dayton, OH 45435  
(513) 873-2403

Degree: Ph. D., Electrical Engr.,  
1975  
Specialty: Digital Control Systems  
Assigned: FDL

Dr. Hemen Ray  
Assistant Professor  
N. C. Agric. & Tech. St. Univ.  
Dept. of Mechanical Engr.  
Greensboro, NC 27411  
(919) 379-7621

Degree: Ph. D., Eng. Mech., 1979  
Specialty: Advanced Composites  
Assigned: FDL



List of Participants (continued: page 18)

Dr. Larry Reeker  
Professor  
Tulane University  
Computer Science Dept.  
New Orleans, LA 70118  
(504) 865-5840

Degree: Ph. D., Comp. Sci., 1974  
Specialty: Computational Linguistics,  
Programming Languages,  
Computer-aided  
Assigned: HRL/Lowry

Dr. David Reynolds  
Assistant Professor  
Wright State University  
Dept. of Engineering  
Dayton, OH 45435  
(513) 873-2403

Degree: Ph. D., Biomedical Eng.,  
1978  
Specialty: Pulmonary Mechanics,  
Biofluid Mechanics  
Assigned: AMRL

Dr. Joseph Saliba  
Assistant Professor  
University of Dayton  
Civil Engr. Dept.  
Dayton, OH 45469  
(513) 299-3847

Degree: Ph. D., Solid Mechanics,  
1983  
Specialty: Structures  
Assigned: AMRL

Dr. Walter Salters  
Associate Professor  
South Carolina State College  
Dept. of Natural Science  
Orangeburg, SC 29115  
(803) 536-7114

Degree: Ph. D., Cell Biology, 1976  
Specialty: Cell and Molecular Biology  
Assigned: SAM

Dr. Lowell Schipper  
Professor  
Bowling Green State University  
Dept. of Psychology  
Bowling Green, OH 43403  
(419) 372-2556

Degree: Ph. D., Psychology, 1953  
Specialty: Statistics, Measurement  
Assigned: HRL/Williams

Dr. Robert Schlegel  
Assistant Professor  
University of Oklahoma  
Dept. of Industrial Engr.  
Norman, OK 73019  
(405) 325-3721

Degree: Ph. D., Ind. Eng., 1980  
Specialty: Human Factors Engineering/  
Ergonomics  
Assigned: AMRL

Dr. Howard Schleier  
Department Chairman  
Norwalk State Tech. College  
Department of Chemistry  
Norwalk, CT 06854  
(203) 838-0601

Degree: M.S., Chem. Eng., 1962  
Specialty: Fluid Mechanics, Heat and  
Mass Transfer,  
Thermodynamics  
Assigned: RPL

List of Participants (continued: page 19)

Dr. James Schneider  
Professor  
University of Dayton  
Physics Department  
Dayton, OH 45469  
(513) 229-2727

Degree: Ph. D., Physics, 1965  
Specialty: Solid State Physics  
Assigned: ML

Dr. Gordon Schrank  
Associate Professor  
St. Cloud State University  
Dept. of Biological Sciences  
St. Cloud, MN 56301  
(612) 255-2036

Degree: Ph. D., Med. Microbiology,  
1974  
Specialty: Electron Microscopy  
Assigned: SAM

Dr. Keith Seitter  
Assistant Professor  
University of Lowell  
Dept. of Earth Sciences  
Lowell, MA 01854  
(617) 452-5000

Degree: Ph. D., Geophysical Sci.,  
1982  
Specialty: Dynamic Meteorology  
Assigned: GL

Dr. Paavo Sepri  
Associate Professor  
University of Oklahoma  
Nuclear Engr. Dept.  
Norman, OK 73019  
(405) 325-5011

Degree: Ph. D., Engr. Science  
1971  
Specialty: Fluid Mechanics and Heat  
Transfer  
Assigned: APL

Dr. Robert Shaw  
Professor  
University of Connecticut  
Dept. of Psychology  
Storrs, CT 06268  
(203) 486-4107

Degree: Ph. D., Psychology, 1965  
Specialty: Perception/Cognition/Math.  
Models/Growth Simulation  
Models  
Assigned: AMRL

Dr. John Sheldon  
Professor  
Florida International Univ.  
Dept. of Physical Sci.  
Miami, FL 33199  
(305) 554-2608

Degree: Ph. D., Engineering,  
1964  
Specialty: Chemical Kinetics  
Assigned: AD

Dr. Harold Sorensen  
Associate Professor  
Washington State University  
Dept. of Civil & Env. Engr.  
Pullman, WA 99164  
(509) 335-5183

Degree: Ph. D., Engineering Mech.,  
1966  
Specialty: Structural Mechanics  
Assigned: WL

List of Participants (continued: page 20)

Dr. Charles Spiteri  
Assistant Professor  
Queensborough Community College  
Dept. of Elect. & Comp. Tech.  
Bayside, NY 11364  
(212) 631-6207

Degree: M.E.E., Elec. Eng., 1980  
Specialty: Computer Communications,  
Local Networks  
Assigned: ESD

Dr. William Squires  
Associate Professor  
Texas Lutheran College  
Dept. of Biology  
Sequin, TX 78155  
(512) 379-4161

Degree: Ph. D., Exercise  
Physiology, 1979  
Specialty: Biology, Physiology  
Assigned: SAM

Dr. Arthur Sterling  
Professor  
Louisiana State University  
Dept. of Chemical Engr.  
Baton Rouge, LA 70803  
(504) 388-1426

Degree: Ph. D., Chemical Engr.,  
1969  
Specialty: Fluid Mechanics, Heat  
Transfer  
Assigned: ESC

Dr. Alexander Stone  
Professor  
University of New Mexico  
Mathematics Dept.  
Albuquerque, NM 87131  
(505) 277-4613

Degree: Ph. D., Mathematics, 1965  
Specialty: Differential Geometry,  
Differential Equations,  
Electromagnetism  
Assigned: SAM

Dr. William Stone  
Assistant Professor  
Meharry Medical College  
Division of Biomedical Sci.  
Nashville, TN 37208  
(615) 327-6506

Degree: Ph. D., Molecular and  
Cellular Biology, 1972  
Specialty: Retinal Biochemistry and  
Lipid Peroxidation  
Assigned: WL

Dr. Jimmy Street  
Associate Professor  
University of Florida  
Soil Science Dept.  
Gainesville, FL 32611  
(904) 392-1951

Degree: Ph. D., Soil Chem., 1976  
Specialty: Soil Pollution  
Assigned: ESC

Dr. John Swetits  
Professor  
Old Dominion University  
Mathematics Department  
Norfolk, VA 23508  
(804) 440-3911

Degree: Ph. D., Mathematics, 1968  
Specialty: Approximation Theory and  
Numerical Analysis  
Assigned: AL

List of Participants (continued: page 21)

Dr. Richard Tankin  
Professor  
Northwestern University  
Mech. & Nuclear Engr. Dept.  
Evanston, IL 60201  
(312) 492-3532

Degree: Ph. D., Mech. Engineering,  
1960  
Specialty: Fluid Mechanics, Heat  
Transfer  
Assigned: APL

Dr. William Thomas  
Assistant Professor  
Meharry Medical College  
Physiology Dept.  
Nashville, TN 37208  
(615) 327-6979

Degree: Ph. D., Biochemistry, 1980  
Specialty: Neurochemistry  
Specialty: SAM

Dr. Ken Tomiyama  
Assistant Professor  
Pennsylvania State University  
Dept. of Electrical Eng.  
University Park, PA 16802  
(814) 865-7667

Degree: Ph. D., System Science,  
1977  
Specialty: Appl. of System Science in  
various Fields  
Assigned: GL

Dr. Albert Tong  
Assistant Professor  
University of Texas  
Dept. of Mechanical Engr.  
Arlington, TX 76019  
(817) 273-2297

Degree: Ph. D., Mechanical Engr.,  
1983  
Specialty: Thermoscience  
Assigned: APL

Dr. Robert Vance  
Assistant Professor  
Ohio State University  
Psychology Department  
Columbus, OH 43210  
(614) 422-0685

Degree: Ph. D., Ind./Organ. Psy.,  
1981  
Specialty: Personnel Psychology,  
Organizational Behavior  
Assigned: HRL/Brooks

Dr. Brian Vick  
Assistant Professor  
Va. Polytechnic Inst. & St. Univ.  
Mech. Engr. Dept.  
Blacksburg, VA 24061  
(703) 961-7596

Degree: Ph. D., Mechanical Engr.,  
1981  
Specialty: Heat Transfer  
Assigned: RPL

Dr. Stephen Wallace  
Assistant Professor  
University of Colorado  
Inst. of Cognitive Science  
Boulder, CO 80309  
(303) 492-8086

Degree: Ph. D., Human Motor  
Behavior, 1976  
Specialty: Human Motor Behavior  
Assigned: HRL/Lowry

List of Participants (continued: page 22)

Dr. Yin-min Wei  
Professor  
Ohio State University  
Computer Science Dept.  
Athen, OH 45701  
(614) 594-6574

Degree: Ph. D., Electrical Engr.,  
1966  
Specialty: Signal Processing  
Assigned: AMRL

Dr. Isaac Weiss  
Assistant Professor  
Wright State University  
Department of Engineering  
Dayton, OH 45435  
(513) 873-3021

Degree: Ph. D., Metallurgy, 1978  
Specialty: Thermomechanical  
Processing, Deformation  
Processing  
Assigned: ML

Martin Werner  
Assistant Professor  
University of Texas  
School of Public Health  
San Antonio, TX 78284  
(512) 691-6845

Degree: Ph. D., Environmental  
Engr., 1982  
Specialty: Hazardous Waste Mgmt.  
Engineering TmT  
Assigned: ESC

Dr. John Wilson  
Associate Professor  
South Carolina State College  
Habilitative Science Dept.  
Orangeburg, SC 29117  
(803) 536-8179

Degree: Ph. D., Speech Science/  
Communications, 1979  
Specialty: Acoustic and Experimental  
Phonetics  
Assigned: RADC

Dr. Krystine Yaworsky  
Associate Professor  
Le Moyne College  
Psychology Department  
Syracuse, NY 13214  
(315) 446-2882

Degree: Ph. D., Cognitive  
Psychology, 1977  
Specialty: Cognitive Processes in  
Perception, Human Learning  
and Memory  
Assigned: HRL/Wright-Patterson

Dr. Chyang Yu  
Assistant Professor  
Wilkes College  
Dept. of Materials Engr.  
Wilkes-Barre, PA 18766  
(717) 824-4651

Degree: Ph. D., Ceramic Eng., 1977  
Specialty: Structural Ceramic  
Materials, Electronic  
Ceramics, Microanalysis  
Assigned: ML

PARTICIPANT LABORATORY ASSIGNMENT (Page 1)

1984 USAF/SCEEE SUMMER FACULTY RESEARCH PROGRAM

AERO PROPULSION LABORATORY

(Wright-Patterson Air Force Base)

- |                     |                        |
|---------------------|------------------------|
| 1. Richard Anderson | 6. Kakkattukuzhy Isaac |
| 2. Huei-huang Chiu  | 7. Richard Miers       |
| 3. Louis Chow       | 8. Paavo Sepri         |
| 4. John Erdei       | 9. Richard Tankin      |
| 5. Dennis Flentge   | 10. Albert Tong        |

AEROSPACE MEDICAL RESEARCH LABORATORY

(Wright-Patterson Air Force Base)

- |                     |                    |
|---------------------|--------------------|
| 1. Gwendolyn Howze  | 5. Joseph Saliba   |
| 2. Robert MacCallum | 6. Robert Schlegel |
| 3. William Norton   | 7. Robert Shaw     |
| 4. David Reynolds   | 8. Yin-Min Wei     |

ARMAMENT DIVISION

(Eglin Air Force Base)

- |                   |                   |
|-------------------|-------------------|
| 1. Myron Calhoun  | 6. Charles Jones  |
| 2. Do Chang       | 7. Robert Kallman |
| 3. Robert Courter | 8. William Pardo  |
| 4. Terrence Dwan  | 9. John Sheldon   |
| 5. Chen-Chi Hsu   |                   |

ARNOLD ENGINEERING DEVELOPMENT CENTER

(Arnold Air Force Station)

1. Gregory Corso
2. Doyle Hasty
3. Madakasira Krishna
4. M. Carr Payne, Jr.

AVIONICS LABORATORY

(Wright-Patterson Air Force Base)

1. Albert Biggs
2. Mike Burlakoff
3. Ronald Greene
4. Paul Griesacker
5. William McCormick
6. James Patterson
7. John Swetits

BUSINESS RESEARCH MANAGEMENT CENTER

(Wright-Patterson Air Force Base)

1. Richard Bernhard
2. Thomas Gullledge, Jr.

PARTICIPANT LABORATORY ASSIGNMENT (Continued: page 2)

ELECTRONICS SYSTEMS DIVISION

(Hanscom Air Force Base)

1. William Perrizo
2. Charles Spiteri

ENGINEERING & SERVICES CENTER

(Tyndall Air Force Base)

- |                           |                    |
|---------------------------|--------------------|
| 1. Annalingam Anandarajah | 5. Datta Naik      |
| 2. Eric Drumm             | 6. Arthur Sterling |
| 3. Paul Hoffman           | 7. Jimmy Street    |
| 4. Yong Kim               | 8. Martin Werner   |

FLIGHT DYNAMICS LABORATORY

(Wright-Patterson Air Force Base)

- |                    |                  |
|--------------------|------------------|
| 1. George Doyle    | 6. Meng Liou     |
| 2. Franklin Eastep | 7. Don Mittleman |
| 3. Albert Havener  | 8. Dale Moses    |
| 4. Mario Innocenti | 9. Kuldip Rattan |
| 5. Walter Jones    | 10. Hemen Ray    |

FRANK J. SEILER RESEARCH LABORATORY

(USAF Academy)

- |                    |                     |
|--------------------|---------------------|
| 1. David Chung     | 5. Hendrik Hameka   |
| 2. Hermann Donnert | 6. Richard Murphy   |
| 3. James Eberhart  | 7. Robert O'Connell |
| 4. Larry Glasgow   |                     |

GEOPHYSICS LABORATORY

(Hanscom Air Force Base)

- |                          |                      |
|--------------------------|----------------------|
| 1. Francesco Bacchialoni | 8. Gabriel Kojoian   |
| 2. John Bahng            | 9. Nabil Lawandy     |
| 3. James Baird           | 10. Bernard McIntyre |
| 4. Alan Bentley          | 11. Martin Patt      |
| 5. Frank Colby           | 12. Keith Seitter    |
| 6. John Cyranski         | 13. Ken Tomiyama     |
| 7. Peter Hierl           |                      |

HUMAN RESOURCES LABORATORY/LRL

(Wright-Patterson Air Force Base)

1. Subramaniam Deivanayagam
2. Emory Ensore, Jr.
3. Sallie Gordon
4. William Kane, Jr.
5. Krystine Yaworsky

HUMAN RESOURCES LABORATORY/OTR

(Williams Air Force Base)

1. Arthur Harriman
2. David Kohfeld
3. Lowell Schipper

PARTICIPANT LABORATORY ASSIGNMENTS (Continued: page 3)

HUMAN RESOURCES LABORATORY/MO

(Brooks Air Force Base)

1. Bruce Feiring
2. Cynthia Ford
3. David Lohman
4. Robert Vance

HUMAN RESOURCES LABORATORY/ID

(Lowry Air Force Base)

1. Larry Reeker
2. Stephen Wallace

LEADERSHIP & MANAGEMENT DEVELOPMENT CENTER

(Maxwell Air Force Base)

1. Louis Buckalew
2. Charles Lardent
3. Kevin Mossholder

LOGISTICS COMMAND

(Wright-Patterson Air Force Base)

1. Kendall Nygard

LOGISTICS MANAGEMENT CENTER

(Gunter Air Force Base)

1. Philip Chong
2. Bruce Janson
3. Evelyn Leggette

MATERIALS LABORATORY

(Wright-Patterson Air Force Base)

- |                          |                     |
|--------------------------|---------------------|
| 1. Charles Drummond, III | 7. Ronald Kline     |
| 2. Delcie Durham         | 8. William Kyros    |
| 3. William Feld          | 9. Larry Ludwick    |
| 4. Thomas Graham         | 10. James Schneider |
| 5. Vijay Gupta           | 11. Isaac Weiss     |
| 6. Vinod Jain            | 12. Chyang Yu       |

OCCUPATIONAL & ENVIRONMENTAL HEALTH LABORATORY

(Brooks Air Force Base)

1. Stephan Nix

ROCKET PROPULSION LABORATORY

(Edwards Air Force Base)

1. Gloria Anderson
2. E. Miller Layton, Jr.
3. Charles Mitchell
4. Howard Schleier
5. Brian Vick



PARTICIPANT LABORATORY ASSIGNMENT (Continued: page 4)

ROME AIR DEVELOPMENT CENTER  
(Griffiss Air Force Base)

- |                  |                        |
|------------------|------------------------|
| 1. Adly Fam      | 6. Robert Jackson, Jr. |
| 2. Basil Gala    | 7. David Lai           |
| 3. Barry Ganapol | 8. Lonnie Ludeman      |
| 4. David Gilliam | 9. John Minor          |
| 5. Brian Holmes  | 10. Johnny Wilson      |

SCHOOL OF AEROSPACE MEDICINE  
(Brooks Air Force Base)

- |                      |                     |
|----------------------|---------------------|
| 1. Deborah Armstrong | 10. James Mrotek    |
| 2. Mukul Banerjee    | 11. Lena Myers      |
| 3. Jeya Chandra      | 12. Boake Plessy    |
| 4. David Cohoon      | 13. Walter Salters  |
| 5. Robert Dorman     | 14. Gordon Schrank  |
| 6. Edward Greco, Jr. | 15. William Squires |
| 7. Kent Knaebel      | 16. William Stone   |
| 8. Raj Krishnan      | 17. William Thomas  |
| 9. Odis McDuff       |                     |

WEAPONS LABORATORY  
(Kirtland Air Force Base)

1. Robert Colclaser, Jr.
2. Frederick Eisler
3. Eddie Fowler
4. Arthur Kovitz
5. Harold Sorensen
6. Alexander Stone

RESEARCH REPORTS  
1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Volume I  
Report  
Number

Title

Research Associate

- |    |                                                                                                                                                               |                              |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| 1  | Centrifuge Modeling of Structural Response During Underground Explosions- a Preliminary Theoretical Feasibility Study                                         | Dr. A. Anandarajah           |
| 2  | New Synthetic Techniques for Advanced Propellant Ingredients: Selective Chemical Transformations and New Structures - Bis-Fluorodinitroethylamino Derivatives | Dr. Gloria L. Anderson       |
| 3  | Infrared Absorption Spectra of Silane and Disilane                                                                                                            | Dr. Richard Anderson         |
| 4  | The Characterization of <sup>14</sup> C-Serotonin Uptake into Cerebellar Glomeruli                                                                            | Dr. Deborah L. Armstrong     |
| 5  | Automatic Controller for Space Experiments                                                                                                                    | Dr. Francesco L. Bacchialoni |
| 6  | Two-Color Refractometry for Astronomical Geodesy                                                                                                              | Dr. John D. R. Bahng         |
| 7  | Long Wavelength infrared Emissions from a Recombining Oxygen Plasma                                                                                           | Dr. James C. Baird           |
| 8  | Development of a High Frequency Lung Ventilation Model for Testing Under Hypobaric Conditions                                                                 | Dr. Mukul R. Banerjee        |
| 9  | Astronomical Observations Using the Imaging Camera of the Large Aperture Infrared Telescope System                                                            | Dr. Alan F. Bentley          |
| 10 | Some Recommendations for Improvements in the Theory and Practice of DoD Incentive Contracting                                                                 | Dr. Richard H. Bernhard      |
| 11 | Simulation of Radar Reception from Terrain and Airborne Targets                                                                                               | Dr. Albert W. Biggs          |

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
12	USAF Spouse Survey: Theoretical Model, Critique, and Revision	Dr. L.W. Buckalew
13	Ada Compiler Evaluation and Validation (E&V) Taxonomy	Mike Burlakoff
14	Software Drivers for the Z-204 Multiport Interface Card	Dr. Myron A. Calhoun
15	The Analysis of a Single Base in an Airlift Operation	Dr. M. Jeya Chandra
16	The Roles of PEI as an Additive in EAK	Dr. Do Ren Chang
17	The Uniqueness of Phase Retrieval from Intensity Measurements	Dr. Huei-huang Chiu
18	Investigating the Potential Application of an "Integrated" Resource Management System to Various Air Force Environments	Dr. Philip S. Chong
19	Low Temperature Expandable Mega-Watt Pulse Power Radiator	Dr. Louis C. Chow
20	Experimental Evaluation and Development of Components for an Infrared Passive Laser Gyro	Dr. David Y. Chung
21	The Electromagnetic Pulse Response of Structures with Frequency Dependent Electrical Properties	Dr. David K. Cohoon
22	Comparison Between Two Atmospheric Boundary Layer Models: Second-Order Turbulence Closure Versus Highly Parameterized, Simplified Physics	Dr. Frank P. Colby, Jr.
23	A Study of the Shiva Star-Type Inductive Pulse Compression System	Dr. R. Gerald Colclaser
24	An Identification of the Human Limitations in the Control Room of Wind Tunnel 4T	Dr. Gregory M. Corso
25	Internal Flow Studies of a Class of Ballistic Launchers	Dr. Robert W. Courter

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
26	Dynamical Spectral Analysis for Nonstationary processes	Dr. John F. Cyranski
27	An Experimental Investigation of Human Torque Strength	Dr. S. Deivanayagam
28	Effects of Nuclear Radiation on the Optical Characteristics of Laser Components	Dr. Hermann J. Donnert
29	Phospholipid Metabolism in a Synaptic Membrane Preparation Isolated from Cerebellar Cortex	Dr. Robert V. Dorman
30	A Review of Computer Simulations for Aircraft-Surface Dynamics	Dr. George R. Doyle, Jr.
31	Constitutive Modelling for Blast Induced Wave Propagation	Dr. E. C. Drumm
32	SiC Reinforced Glass-Ceramic Composites	Dr. Charles H. Drummond
33	Computer Simulation of TI-6AL-4V and Rene' 95 Disks for the T-700 Engine	Dr. Delcie R. Durham
34	Graphics Generation and Image Enhancement and Restoration Techniques	Dr. Terrence E. Dwan
35	Structural Modifications to Enhance the Active Vibration Control of Large Flexible Structures	Dr. Franklin E. Eastep
36	Wetting Behavior of Imidazolium-Containing Room-Temperature Molten Salts	Dr. James G. Eberhart
37	The Emittance of Particle and Laser Beams and Measurement of the Angle Between Crossed Laser and Particle Beams to High Precision	Dr. Frederick R. Eisler
38	A Generic Logistics Model for Evaluating Operational Readiness of a Weapon System	Dr. E. Emory Enscoe, Jr.

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
39	Characterization of Turbulence Through Methods of Field Theory	Dr. John E. Erdei
40	Issues in High Throughput Signal Processing	Dr. Adly T. Fam
41	Upgrade of Policy Specifying	Dr. Bruce Feiring
42	New Phenoxy Substituted Dianhydrides	Dr. William A. Feld
43	Cyclic Voltammetric and Cobra Analysis of Synthetic Lubricant Degradation	Dr. Dennis R. Flentge
44	Gender Differences on Subtests of the ASVAB and the Relationship Between ASVAB Subtest Scores and Psychological/Social Variables for Males and Females	Dr. Cynthia A. Ford
45	Communications Network Simulation Topics	Dr. Eddie R. Fowler
46	Nonparametric CFAR Detection of MTI Radar Signals in Heavy Clutter	Dr. Basil E. Gala
47	Benchmark Solutions for the Spencer-Lewis Equation Describing Electron Transport in an Infinite Medium	Dr. Barry D. Ganapol
48	Electromagnetic Waves in a Disturbed Atmosphere Environment	Dr. David S. Gilliam
49	Alternative Computational Methods for Separated Flows on Pitched Airfoils	Dr. Larry A. Glasgow
Volume II		
50	Manual and Computer-Aided Sequential Diagnostic Inference	Dr. Sallie E. Gordon
51	Angle Resolved Ion-Scattering Spectroscopy—a Feasibility Study	Dr. Thomas P. Graham
52	Electrogastrogram and its Effectiveness in Evaluation of Motion Sickness	Dr. Edward C. Greco, Jr.

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
53	Far-Infrared Absorption Profiles for Shallow Donors in GaAs-GaAlAs Quantum Well Structures	Dr. Ronald L. Greene
54	Digital Signal Processing Approaches for Analysis and Evaluation of Communication Systems	Dr. Paul B. Griesacker
55	Production Rate Variations Cost Models	Dr. Thomas R. Gullledge, Jr.
56	Thermal Stability Characteristics of Silahydrocarbons	Dr. Vijay K. Gupta
57	Calculations of Electron Spin Resonance Coupling Constants	Dr. Hendrik F. Hameka
58	Effects of Pyridostigmine on Performance of Mission-ready Pilots in the OT Simulation Facility	Dr. Arthur E. Harriman
59	Propulsion Facility Planning for Test Information Productivity Improvement with Emphasis on Data Measurement Uncertainty in the Engine Test Facility	Dr. Doyle E. Hasty
60	A Study on Point Diffraction Interferometry Plus Holographic Measurements of a Turbulent Boundary Layer on a Roughened Wind Tunnel Wall	Dr. A. George Havener
61	Effects of Temperature and Reactant Solvation Upon the Rates of Gas-Phase Ion-Molecule Reactions	Dr. Peter M. Hierl
62	A Monte Carlo Sampling of BDR Times	Dr. Paul C. Hoffman
63	Bismuth Silicon Oxide: Sample Variability Investigated with Thermally Stimulated Conductivity and Thermoluminescence	Dr. Brian W. Holmes

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
64	Comparison of Periosteums from Femur and Vertebral Bone	Dr. Gwendolyn B. Howze
65	On a Thin-Layer Navier-Stokes Code and Transonic Projectile Aerodynamics	Dr. Chen-Chi Hsu
66	Effect of Display Dynamics in Manual Control Tasks	Dr. Mario Innocenti
67	Computational Study of Ramjet Compustor Flowfields	Dr. Kakkattukuzhy M. Isaac
68	Numerical Characterization of Microstrip Discontinuities on Thick Substrates	Dr. Robert W. Jackson, Jr.
69	Advanced Physical Modeling/ Wedge Test Development	Dr. Vinod K. Jain
70	Development of a Vehicle Fleet Analysis System	Dr. Bruce N. Janson
71	Photographic Emulsions for Preparation of Holographic Filters	Dr. Charles R. Jones
72	Fracture Behavior of Cross-Ply Graphite/Epoxy Composite Laminates	Dr. Walter F. Jones
73	The Optimal Construction of Synthetic Discriminant Functions for Optical Matched Filters	Robert R. Kallman
74	Stress and Aircraft Maintenance Performance in a Combat Environment	Dr. William D. Kane, Jr.
75	Finite Element Analysis of Centrifuged Concrete Culverts	Dr. Yong S. Kim
76	Acoustic Emission in Composites	Dr. Ronald A. Kline
77	Process Configuration Alternatives for Separation of Gas Mixtures by Pressure Swing Adsorption	Dr. Kent S. Knaebel
78	The Distributional Analysis of Contrast Sensitivity Measures	Dr. David L. Kohfeld

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
79	The Relation Between Hard X-Ray Bursts and Type II Radio Emission	Dr. Gabriel Kojoian
80	Thermal Layer Development with Energy Flux at the Ground/Air Interface	Dr. Arthur A. Kovitz
81	Computational Fluid Dynamics Grids-Flow Properties Inter- pretation Algorithm	Dr. Madakasira G. Krishna
82	Development of an Optical Multichannel Analyzer System	Dr. Raj. M. Krishnan
83	Interlaminar Shear Testing of Carbon/Carbon Composites	Dr. William Kyros
84	Two-Dimensional Median for Image Preprocessing in Machine Recognition	Dr. David C. Lai
85	Psychological Correlates of Physiological Indicators of Stress-Related Disorders: A Search for Structure and Relatedness	Dr. Charles L. Lardent, Jr.
86	Optical Bistability with Liquid Media: Experimental Studies and Theoretical Predictions	Dr. N.M. Lawandy
87	Survey of High-Energy Molecular Systems	Dr. E. Miller Layton, Jr.
88	Rewrite of AFM 28-345	Dr. Evelyn J. Leggette
89	Numerical Simulation of a Supersonic Inlet Flow	Dr. Meng-Sing Liou
90	Information Processing Analysis of Spatial Synthesis and of the Relationship Between Learning and Intelligence	Dr. David F. Lohman
91	Suboptimum Extrapolation for Spectral Estimation	Dr. Lonnie C. Ludeman



<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
92	Silane-Treated Silica Fillers for Elastomer Reinforcement	Dr. Larry M. Ludwick
93	The Application of Structural Equation Modeling to Experimental Data	Dr. Robert MacCallum
94	The use of the Instantaneous Frequency Transient in the Design and Optimization of the Channelized Receiver and Instantaneous Frequency Measurement (IFM) Versions of the Passive EW Receiver	Dr. William S. McCormick
95	Techniques for Ultra-Short Pulses in Nd:YAG Lasers	Dr. Odis P. McDuff
96	Plasma Generation and Diagnostics for Ionospheric Plasma Simulation	Dr. Bernard McIntyre
97	Spectroscopic Studies of Thyratron Discharges	Dr. Richard E. Miers
98	On Interfacing Logic Programming Systems and Relational Databases	Dr. John T. Minor
Volume III		
99	Evaluation of Models for Liquid Propellant Rocket Combustion Instability	Dr. Charles E. Mitchell
100	Large Space Structure Dynamic Testing	Dr. Don Mittleman
101	Acquisition of Wind Tunnel Wall Pressure Distributions for use in Developing A 3-D Transonic Wall Correction Code	Dr. Dale F. Moses
102	Leadership Effects as Measured by the Organizational Assessment Package: A Multi- level Perspective	Dr. Kevin W. Mossholder
103	Raman Spectroscopy of Unstimulated and Stimulated Cultured Normal and Neoplastic Human or Mammalian Cells	Dr. James J. Mrotek

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
104	Preliminary Monte Carlo Studies of the Structure of Molten Salts	Dr. R. D. Murphy
105	Military Family Stress	Dr. Lena Wright Myers
106	Air Oxidation of Hydrazine - A Kinetic Study	Dr. Datta V. Naik
107	Conceptual Design of the USAF Installation Restoration Program Information Management System	Dr. Stephan J. Nix
108	The Cytotoxic Effects of Trimethylpentane on Rat Renal Tissue	Dr. William N. Norton
109	Computer-Based Optimization Algorithms for LOGAIR Cargo Allocation	Dr. Kendall E. Nygard
110	Laser Damage Studies in Purified and Plasticized Polyakylmethacrylates	Dr. Robert M. O'Connell
111	Experimental Physics Aspects of the AFATL Railgun Effort	Dr. William B. Pardo
112	Analysis of the Validity of Barnes Transmissometer Data	Dr. Martin A. Patt
113	Permanent Periodic Magnets and the Reproducibility of Traveling Wave Tubes	Dr. James D. Patterson
114	Operator Activities in Wind Tunnel 4T	Dr. M. Carr Payne, Jr.
115	Future Tactical Air Control System Database Design	Dr. William Perrizo
116	Raman Spectroscopy of Glycosaminoglycans from Bovine Cornea	Dr. Boake L. Plessy
117	Study of Control Mixer Concept for Reconfigurable Flight Control System	Dr. Kuldip S. Rattan

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
118	Analylsis of Armor Bracketry	Dr. Hemen Ray
119	Artificial Intelligence and Computational Linguistics Research	Dr. Larry H. Reeker
120	Mathematical Modeling of the Human Cardiopulmonary System	Dr. David B. Reynolds
121	Nonlinear Modeling of Seat Cushions	Dr. Joseph E. Saliba
122	The Relationship of Fibrinogen and Plasma Lipids to Bubble Formation at the Air Interface During Decompression Sickness	Dr. Walter L. Salters
123	N/A	Dr. Lowell Schipper
124	Evaluation of Training Performance for the USAF Criterion Task Set	Dr. Robert E. Schlegel
125	Regenerative Heat Transfer in an LSSCDS Engine	Dr. Howard Schleier
126	Raman Spectroscopy Studies of Extrinsic P-Type Silicon	Dr. James Schneider
127	Bacteriologic Techniques for the Isolation and Identification of Legionellae	Dr. Gordon D. Schrank
128	A Three-Dimensional Radiation Boundary Condition for Mesoscale Numerical Models	Dr. Keith L. Seitter
129	Calculation of Enhanced Heating in Turbulent Boundary Layers Influenced by Free Stream Turbulence	Dr. Paavo Sepri
130	An Adjoint Systems Approach to Learning and Transfer of Training	Dr. Robert E. Shaw
131	The Theory of Self-Heating Phenomena in Explosives with Applications to EAK	Dr. John W. Sheldon

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
132	The Development of Computational Efficiencies in Continuum Finite Element Codes using Matrix Difference Equations	Dr. Harold Sorensen
133	Base Communications Architecture Security Issues	Dr. Charles J. Spiteri
134	Cardiovascular Responses of High- and Low-Fit Men to Head-Down Rest Followed by Orthostasis and Exercise	Dr. William Squires
135	Recommendations on Combustion Research at Tyndall Air Force Base, Florida	Dr. Arthur M. Sterling
136	Electromagnetic Lens Design Techniques	Dr. Alexander P. Stone
137	The Role of Antioxidant Nutrients in Preventing Hyperbaric Oxygen Damage to the Retina	Dr. William L. Stone
138	Naphthalene Adsorption by Florida Soils	Dr. Jimmy J. Street
139	An Optimal Trajectory Problem	Dr. John J. Swetits
140	The Role of Vortex Shedding in a Bluff-Body Combustor	Dr. Richard S. Tankin
141	An Investigation of Acetylcholine as a Neurotransmitter of Cerebellar Mossy Fibers	Dr. William E. Thomas
142	Unified Real Part of Susceptibility over Millimeter through Infrared Region	Dr. Ken Tomiyama
143	Numerical Modeling of Multi-phase Turbulent Recirculating Flows in Sudden-Expansion Ramjet Geometry	Dr. Albert Y. Tong

<u>Report Number</u>	<u>Title</u>	<u>Research Associate</u>
144	Development of Three Covariance Structure Models for Analysis of Performance Measurement Project Data	Dr. Robert J. Vance
145	Laminarization in Highly Accelerated Flow	Dr. Brian Vick
146	Development of Computer Assisted Instruction in Basic Electronic Trouble Shooting	Dr. Stephen A. Wallace
147	The Development of a Computerized System for Management Information and Inter-Office Communication	Dr. Yin-min Wei
148	The Processing Window for the Near Beta Ti-10V-2Fe-3Al Alloy	Dr. Isaac Weiss
149	Effects of Humidity on Gaseous Phase Adsorption of Trichloroethylene by Activated Carbon	Dr. Martin D. Werner
150	A Comparative Analysis of Whispered and Normally Phonated Speech using an LPC-10 Vocoder	Dr. Johnny R. Wilson
151	Cognitive Factors in Computer- Aided Fault Diagnosis	Dr. Krystine B. Yaworsky
152	Characterization of Ceramic- Ceramic Composites	Dr. Chyang John Yu

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

MANUAL AND COMPUTER-AIDED SEQUENTIAL DIAGNOSTIC INFERENCE

Prepared by:	Dr. Sallie E. Gordon
Academic Rank:	Assistant Professor
Department and University:	Department of Psychology University of Idaho
Research Location:	Air Force Human Resources Laboratory, Logistics and Human Factors Division, Ground Operations Branch
USAF Research:	Rosemarie J. Preidis
Date:	August 20, 1984
Contract No:	F49620-82-C-0035

MANUAL AND COMPUTER-AIDED SEQUENTIAL

DIAGNOSTIC INFERENCE

by

Sallie E. Gordon

ABSTRACT

It is becoming increasingly obvious that computerized automation can be a useful aid for a wide variety of positions in the Command and Control network, where many of the tasks involve situational assessment or "diagnostic inference". To optimally combine human talent and computer-aiding systems we must know how the human operator performs the task unaided (and under what circumstances), what subtasks can be allocated to the machine, and what variables affect operator acceptance of the aiding system. This paper presents a theoretical model of the human performance of a diagnostic inference task when unaided by machine including the variables affecting those inference processes, and a preliminary model of how a computer-aiding system might be expected to fit into the diagnostic system.

### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for making possible a very interesting and rewarding summer at the Air Force Human Resources Laboratory, Wright-Patterson AFB, Ohio. She would like to express sincere appreciation to the laboratory, in particular the Ground Operations branch for an exceptional working environment.

Finally, she would like to thank Ms. Rosemarie Preidis for her enormous support and collaboration in a difficult undertaking, Bertram Cream for his insightful administrative guidance, Larry Reed for many helpful discussions, and the other members of the branch for their professional support and friendship.



## I. INTRODUCTION:

It is becoming increasingly obvious that computerized automation can be a useful aid in a wide variety of positions in the armed services. This is especially true in the world of Command and Control (C<sup>2</sup>) where much of the work involves complex situational assessment or "diagnostic inference". As technological complexities increase, the human operator will have a more difficult time trying to understand, integrate, and utilize the information made available to him. In contrast to man's limited cognitive capacities and well-documented biases<sup>1,2</sup> a computer can utilize and aggregate large volumes of information using pre-determined optimal strategies that are most appropriate for the situation at hand. It is no longer a question of whether computer aiding will be used, but how it will be used.

Just as there are problems inherent in using a completely "manual" system to perform a function, there are also problems in using a completely "automated" system. These problems have been discussed in length elsewhere,<sup>2,3</sup> but let it suffice to say that at the current time expert systems are not sufficiently advanced to make automated systems infallible or able to deal with the multitude of unforeseen occurrences that are likely in the C<sup>2</sup> environment.

Since neither man nor machine are solely capable of performing situational assessment functions, the solution lies in using both together and relying on the strengths of each (hopefully also minimizing the weaknesses of each). To integrate man and machine successfully for a given task, one must understand how the human perceives and performs the task, how the machine can be programmed to perform the task (or parts of the task) and then analyze the best way to fit the two together. In the procurement cycle, a common method for developing a computerized aiding system is to intuitively develop a software system that seems like it could do the job. Little attention is paid to analysis of the entire task and what subtasks could be best performed by the man and which are best left to the machine (a few exceptions do exist). Consideration is usually not given beforehand to how the operator will react to the aiding system, nor to what variables will lead to his acceptance or rejection of the new system. Instead, an automated system is designed

and a prototype built. Any modifications necessary to make the system compatible with the operator are usually done after this point. This leads to only those changes that seem absolutely necessary and the result is an overall man-machine system that is much less effective and efficient than what could have been achieved.

Part of the problem outlined above results from our inadequate knowledge concerning three vital questions: (1) How does the human operator perform the assessment task when unaided by automation? (2) What subtasks are best performed by the operator and what subtasks are best performed by the automation device? (3) What factors determine operator acceptance and use of the automated system? The first question (how the operator performs the task) may seem unnecessary to some. However, we do need this information because it directly affects the answers to the second and third questions. That is, if we know how the operator performs the task (not how he differs from some theoretically optimal strategy) we can determine specifically what capabilities he has that we want to preserve in determining the optimal man-machine subtask allocation. In addition, one can argue that how the operator performs the task will largely affect his acceptance of the automation. If the automation is extremely different from or incompatible with his way of perceiving and accomplishing the task, then he will be less likely to accept and use that automation.

The theories and methodologies of cognitive psychology can be brought to bear on this problem. By mapping out the cognitive processes or strategies that are used by the perceiver under various situational constraints, we can then measure how those processes change as a function of providing a computer aiding system.

## II. OBJECTIVES:

The primary objective of this effort was to develop a predictive model of the diagnostic inference task, and how that task would be affected by implementation of an automated system. This objective included the following specific goals:

- (1) Develop a descriptive model of the inference task that characterized inference tasks in the  $C^2$  system and would be amenable

to laboratory research.

- (2) Determine appropriate techniques for measuring process and performance in the inference task.
- (3) Determine a preliminary set of independent variables expected to affect the inference process (and performance).
- (4) Develop a predictive model of the effects of automation on the operator's inference processes.

Accomplishing these goals would serve two purposes: provide guidance to researchers at AF Human Resources Laboratory concerning variables of critical interest in related field research, and provide a framework for follow-on laboratory research designed to answer some of the questions outlined earlier.

### III. DESCRIPTION OF THE INFERENCE TASK:

Diagnostic inference will be defined as the process of using available cues to determine the underlying or "unseen" cause of those cues. An example is medical diagnosis where the doctor must infer a disease causing some set of symptoms. If the available cues are very informative, the inference will be accurate and made with a high degree of confidence. However, it is often the case that the cues do not convey enough information and the inference task takes place amid psychological uncertainty.

In the past, most research addressing this type of task assumed a "single-stage" process, where the perceiver received the cues and somehow aggregated or operated upon the information and derived a judgment. This was a popular view for some time, partially because it was amenable to laboratory experimentation and formal mathematical description and analysis.<sup>4</sup> Two approaches were common; the first was to develop a formal mathematical model (such as Bayes Theorem) to specify optimal performance and then fit that model to data obtained with human subjects.<sup>5</sup> The other approach was to use linear regression models to assess how the subjects were combining or utilizing the cues in generating the inference.<sup>6,7</sup> Research questions addressed in these studies included such topics as, what cues are predominantly utilized by people<sup>7</sup>, how many dimensions or cues are used for various tasks and are they the

7

same as in the "real world" <sup>8</sup>, and whether experts cluster or weigh cues in a similar fashion <sup>9,10</sup>.

The appropriateness of these models has been debated recently and this issue will not be discussed in length. However, two points will be reviewed. The first is the criticism that most laboratory inference tasks involve simultaneous and orthogonally manipulated cues. This cue independence is seen as being highly artificial and unrealistic <sup>11</sup>. Since humans develop cognitive skills to deal with a real, complex world, it is not surprising that they perform "suboptimally" on these inference tasks where no intercue correlations are preserved.

The second criticism with these approaches suggests that we should treat the inference task not as a single-stage process but as a multiple-stage process. <sup>12</sup> This is not to say that the reception vs. the integration of cues are different stages, but that the acquisition of cues or characteristics takes place over time and that this process should be reflected in the theoretical models.

In line with these criticisms and recent views in cognitive psychology, it will be assumed that the inference task of interest takes place in a complex situation where the perceiver must sequentially seek information to make the inference judgment. In addition, that information is typically incomplete and varies in its diagnosticity. The perceiver starts with one or two cues and then searches for others to either confirm hypothesized causes or suggest new ones. The inference process is viewed as a "constructive" process, much like building a jigsaw puzzle. One does not need all of the pieces to be able to infer the nature of the picture, instead the ability to draw the inference will depend on the combined information provided by the pieces put together.

In psychological terms, the perceiver uses both Conceptually-Driven processing (where the hypothesized cause suggests cues to seek) and Data-Driven processing (where cues suggest plausible hypotheses). The cyclic procedure continues until the perceiver exceeds some certainty criterion that he/she knows the identity of the cause. In some cases, a lack of information will prevent that criterion from being reached at all.

#### IV. MEASURING PROCESS AND PERFORMANCE:

It was suggested that the inference judgment is constructed over time as information is acquired. This implies that it is important to measure the process by which the perceiver is coming to a conclusion as well as to measure performance per se. Each of these issues will be addressed in turn.

##### A. PROCESS

Several methodologies for measuring judgment or decision "process" have been suggested. Payne<sup>13</sup> is a predominant supporter of two of these methods known as process tracing. The first method is a class of measurement techniques where the subjects' information acquisition is monitored. The subject must view or select information in such a way that can easily be observed and recorded. Data is obtained concerning what cues the subject samples, in what order, how many are sampled, and the amount of time for the cue sampling.

The other method of process tracing is the collection of verbal protocol. In this technique, the subject is simply asked to "think out loud" as he/she performs the task. Although this type of data can give us insight into the subject's strategies, it cannot be assumed that the subject will always verbalize the cognitive processes as they occur.

After assessing the various process measurement techniques, a method was decided upon which seems most suitable to an inference task. The inference is actually a classification task where the perceiver must choose between class A, class B, class C, and so on (also possibly "none of the above"). The process measure being suggested consists of two aspects:

- (1) Allowing the subjects to acquire whichever cues they desire until they feel reasonably confident in their choice (this is similar to the previously described information acquisition measure), and

- (2) Asking the subject after each cue acquisition to give the hypothesized cause(s) along with a subjective certainty rating (i.e., 1=not certain at all, 7=extremely certain). An example of these measures will be presented shortly.

##### B. PERFORMANCE

In addition to studying cognitive processes or strategies engaged in by the perceiver, it is informative to determine how well the operator is able to infer the cause of the cues. The most appealing measures are suggested by Signal Detection Theory (SDT) because it allows for separate measurement of discrimination capabilities and subjective bias.<sup>14</sup> However, SDT cannot at this time be applied to more than a two-category (Signal-Noise) task. Swets and Pickett<sup>15</sup> discuss the problem of multiple category discrimination and suggest using % Correct as a reasonable solution. This is justified because the inference task for multiple causes is actually conceptually similar to a forced choice task. It was therefore determined that % Correct could be used as a performance measure in the diagnostic inference task.

To summarize the inference task and associated measures of process and performance:

- the subject is given an initial cue
- the subject verbalizes one or more hypotheses along with a certainty rating
- the subject samples a new cue of his/her choice
- the subject verbalizes revised hypothesis(es) along with new certainty rating

:

(continues until subject exceeds some subjective certainty criterion)

An example of data collected from a subject is given below:

	<u>class A</u>	<u>class B</u>	<u>Neither</u>
Cue 1	1		
Cue 2	2	2	
Cue 3		4	
Cue 4		7	

It can be seen that after the first cue sampling, the subject hypothesized class A as the cause. However, after receiving additional cues, the hypothesized cause was switched to class B with increasing certainty.

To determine performance (% Correct) from the data, a cutoff point

must be chosen for the certainty scale. For example, correct might be arbitrarily defined as "four and above". If the correct answer for the example was class B, the subject would be scored as "correct" on only the last observation given above. The advantage of this method is that accuracy of the inference can be assessed without the subject having to make a strictly Yes-No decision (this is a way of "setting" the subject's decision criterion). With this method of data collection, it is possible to measure a variety of process and performance variables as the subject samples the information:

- (1) information sampled
- (2) hypotheses generated
- (3) hypothesis transition (when do they give up and generate another)
- (4) subjective certainty criteria
- (5) accuracy

#### C. FEASIBILITY STUDY

A pilot study was conducted during the Summer Fellowship period to determine the appropriateness of the above variables. Eight subjects (students at Wright State University, Dayton, Ohio) received course credit for participating in the study.

Subjects were first given fictitious names of two diseases along with eight case studies for each disease. The case studies were described in terms of patient initials, occupation, and a set of symptoms that varied in number from three to seven. There were eight total symptom dimensions (such as blood pressure, weight loss, etc.). Four of the dimensions were strongly associated with disease A, and the other four were strongly associated with disease B. However, there was some overlap of symptoms across diseases. After subjects reported that they were familiar with the disease characteristics (approximately 15 minutes), they were given 32 new case studies to diagnose. For each one, they were presented with a 3x5 index card with initials, occupation, and one symptom. They were also presented with seven other cards with a symptom dimension (e.g., "blood pressure") labelled on one side. They were told to turn any card over that they wished, and after each one to tell the observer their hypothesis and certainty rating. Data was collected corresponding to the example shown previously.

Based on the theoretical considerations outlined earlier, two effects were expected:

- (1) The certainty ratings would start low and slowly increase until some criterion was reached - at this point subjects would discontinue information acquisition.
- (2) The % Correct would be a positive function of the number of symptoms sampled on a given trial.

Data relevant to the first hypothesis are presented in Figure 1. It can be seen that certainty ratings did in fact start relatively low and increase over cue samples; the rate of increase was highest for trials where subjects only sampled 4 cues, and lowest for trials where subjects sampled 8 cues. Notice that for trials where all symptoms were utilized, the certainty rating never reached a level equal to the other trials. That is, a subjective certainty criterion of approximately 5 was exceeded in all but those trials where subjects simply did not have enough information to generate an inference with any degree of confidence.

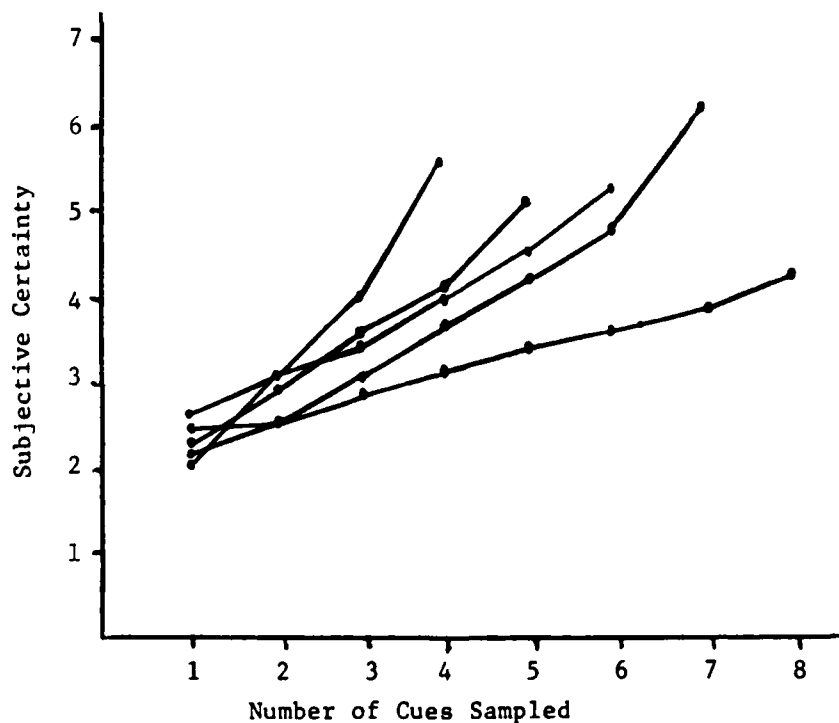


Figure 1. Subjective Certainty as a Function of Cues Sampled



This pattern in the Certainty data is consistent with the Accuracy (% Correct) data shown in Figure 2. A liberal cutoff was arbitrarily chosen so that "three and above" for the Certainty rating was considered a "correct" answer. Even still, the scores were not remarkably high (mean % correct for all trials was .66). This would indicate that subjects were not able to completely learn the correct structure of the disease-symptom associations. Performance varied widely from subject to subject, with a range of .50 to .97 for the eight subjects.

Figure 2 shows that subjects' perceived uncertainty had some basis, performance decreased as the subjects utilized a greater total number of symptoms. Interestingly, both data sets show the same pattern of an increase in accuracy and subjective certainty for trials where seven cues were sampled. Additional studies will be conducted to determine whether this effect still holds for larger sets of data.

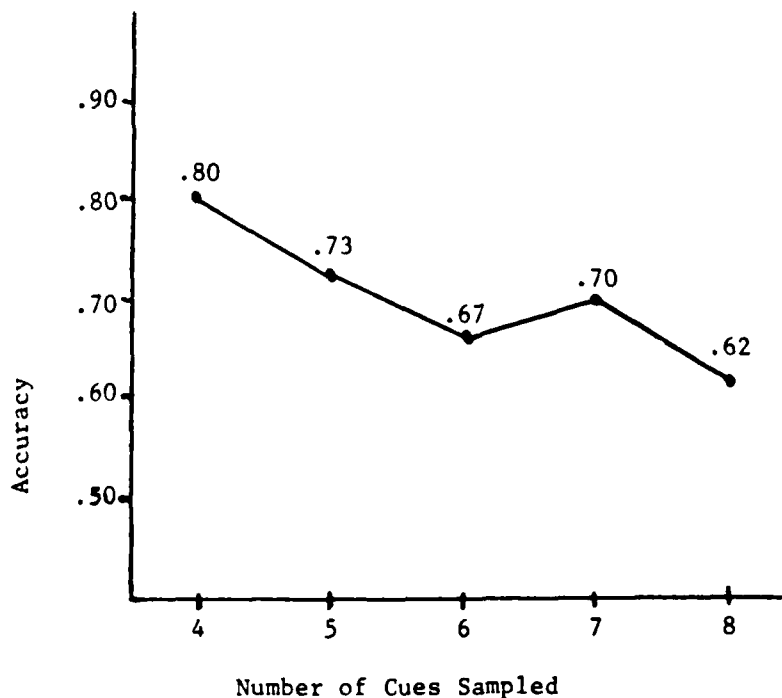


Figure 2. Accuracy as a Function of Cues Sampled.

It can be seen from the data collected thus far, that the process and performance measures are appropriate for the task and yield a rich variety of information concerning the perceiver's strategy. (It should be noted that additional more fine-grained analyses are planned.)

#### V. VARIABLES AFFECTING THE INFERENCE PROCESS:

Inference process and performance are each affected by various situational constraints. These include characteristics of the perceiver, of the task, and of the situational environment from moment to moment. A real world illustration of this complex situation will be described as a way of introducing situational constraints which affect process and performance variables.

An important inference function within the C<sup>2</sup> network is that of the radar operator (and/or officer) who must determine the identity of aircraft showing up on the radar scope. This person receives a "track" on his scope and must "infer" the identity of the aircraft using auxillary pieces of information or "cues". These pieces of information include Flight Plan Data, Special Codes emitted by the aircraft (friendly aircraft), speed, heading, electronic emissions, intelligence data, and possibly visual identification information. Some of this information will be quite diagnostic (i.e., the Code), while others will not be particularly diagnostic (speed or heading). Given enough time, the operator could identify almost any track - if nothing else he could send someone up to look at the aircraft. The problem is based on the fact that the operator officially has two minutes maximum to identify the aircraft. In addition, sending someone up for visual identification is very costly. In wartime conditions, the operator will have to identify many tracks in a very short period of time.

A preliminary list of variables was developed that are considered most important in determining the operators performance (as defined by accuracy). These situational constraints are shown in Figure 3; for example, "time stress" is a variable that will negatively impact performance. Time stress will be determined by the number of tracks the operator has to identify, how far away the aircraft are from important locations, and the alert status at the time (i.e., white, yellow, or red alert).

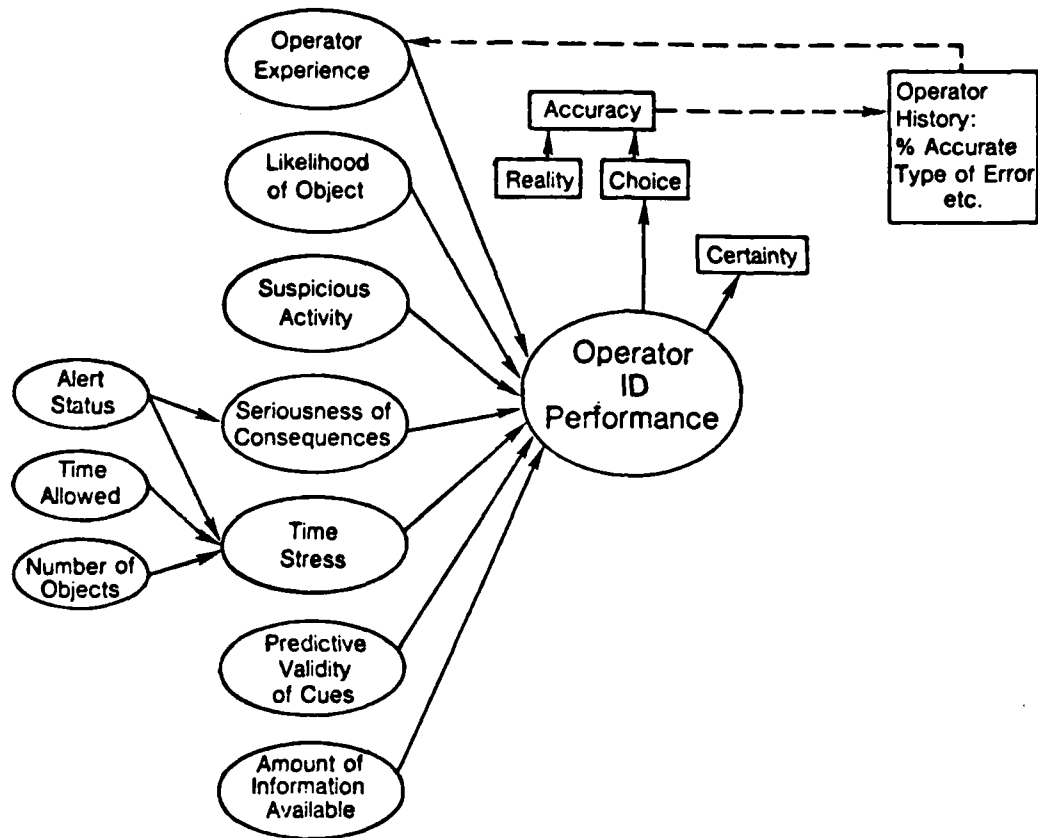


Figure 3. Variables Affecting Operator Performance

The same variables are assumed to also impact the process of generating an inference. Time stress will result in fewer cues considered and a lower subjective certainty criterion (point where the perceiver is willing to stop collecting data). Time stress is also expected to decrease the number of hypotheses generated and considered during the inference process.

Suspiciousness of the aircraft travel will increase the likelihood of an enemy aircraft being hypothesized, resulting in cues being sampled which will confirm (or disconfirm) that hypothesis. Diagnosticity or predictive validity of the cues will result in a need to sample fewer cues and a higher subjective certainty concerning the inference. Finally, operator experience is assumed to enhance the inference performance, but it is not totally clear how that variable will impact the process.

## VI. IMPACT OF AUTOMATION:

Before discussing the impact of implementing a computer aiding system, it is necessary to define the nature of the automated system. Several researchers have developed taxonomies of types of automation,<sup>16</sup> most are admittedly imprecise. For the present purposes, a continuum will be assumed with a completely manual method of task accomplishment at one end and complete automation at the other end of the continuum. An example of this type of continuum is shown in Figure 4. One of the more prevalent types of automation is given on the far right of the scale, that is, the computer completely performs the task and provides an "answer" to the operator who then decides if he/she wants to believe and utilize that answer. This type of automation is currently being planned for the radar identification task described earlier. A predictive model can be outlined at this point describing the impact of implementing this type of automation. First, we can say that most of the variables expected to affect the human perceiver will probably not affect the performance of the automated system. Thus under conditions of stress, the human performance will deteriorate while the machine performance will not. The only variables expected to affect the machine performance will be the characteristics of the cues themselves (see Figure 5). This puts the operator in the place of deciding whether to "trust" the machine knowing that the machine can perform the task more quickly and objectively in times of duress. His decision to use the answer provided by the system will depend on how much time he has, the seriousness of the consequences, and the nature of the information (i.e., is it suspicious activity?). In addition, his decision to use the answer provided will strongly depend on his feelings of his own ability vs. the history of the machine reliability and accuracy. If the machine has a relatively low "hit rate", he will be more inclined to consider it worthless and go mostly on his own judgment.

An interesting point which should be considered is the possibility that the operator might treat the answer provided by the machine as just "another cue". Rather than determining his inference based on the standard set of cues and then comparing this with the automated choice, the operator may just use that choice during the inference process treating it as he does all of the other pieces of information.

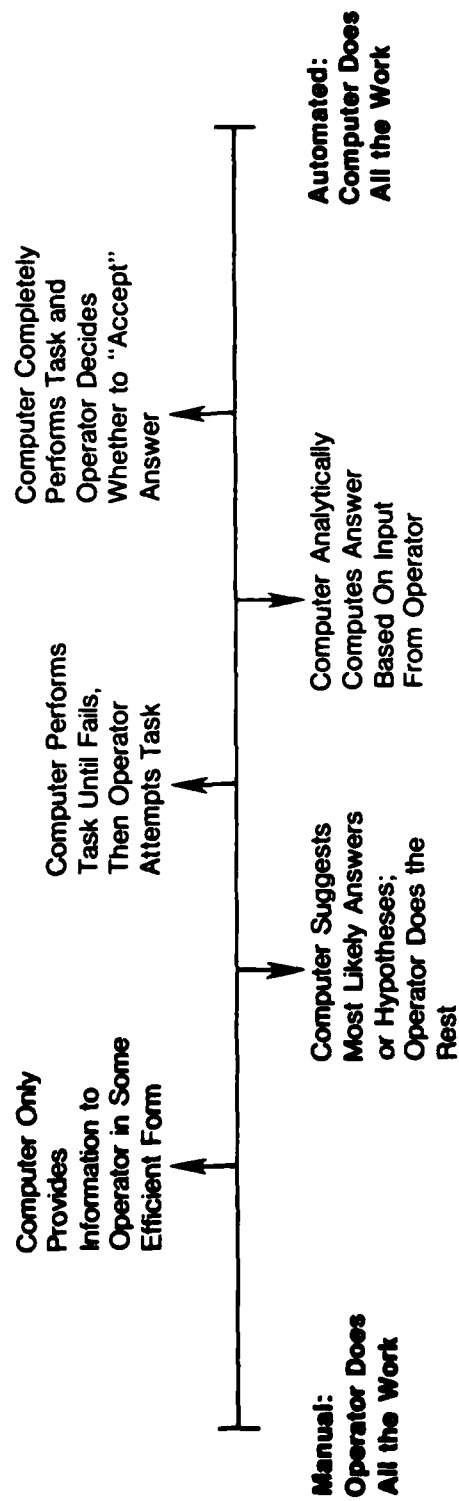


Figure 4. Degrees of Automation

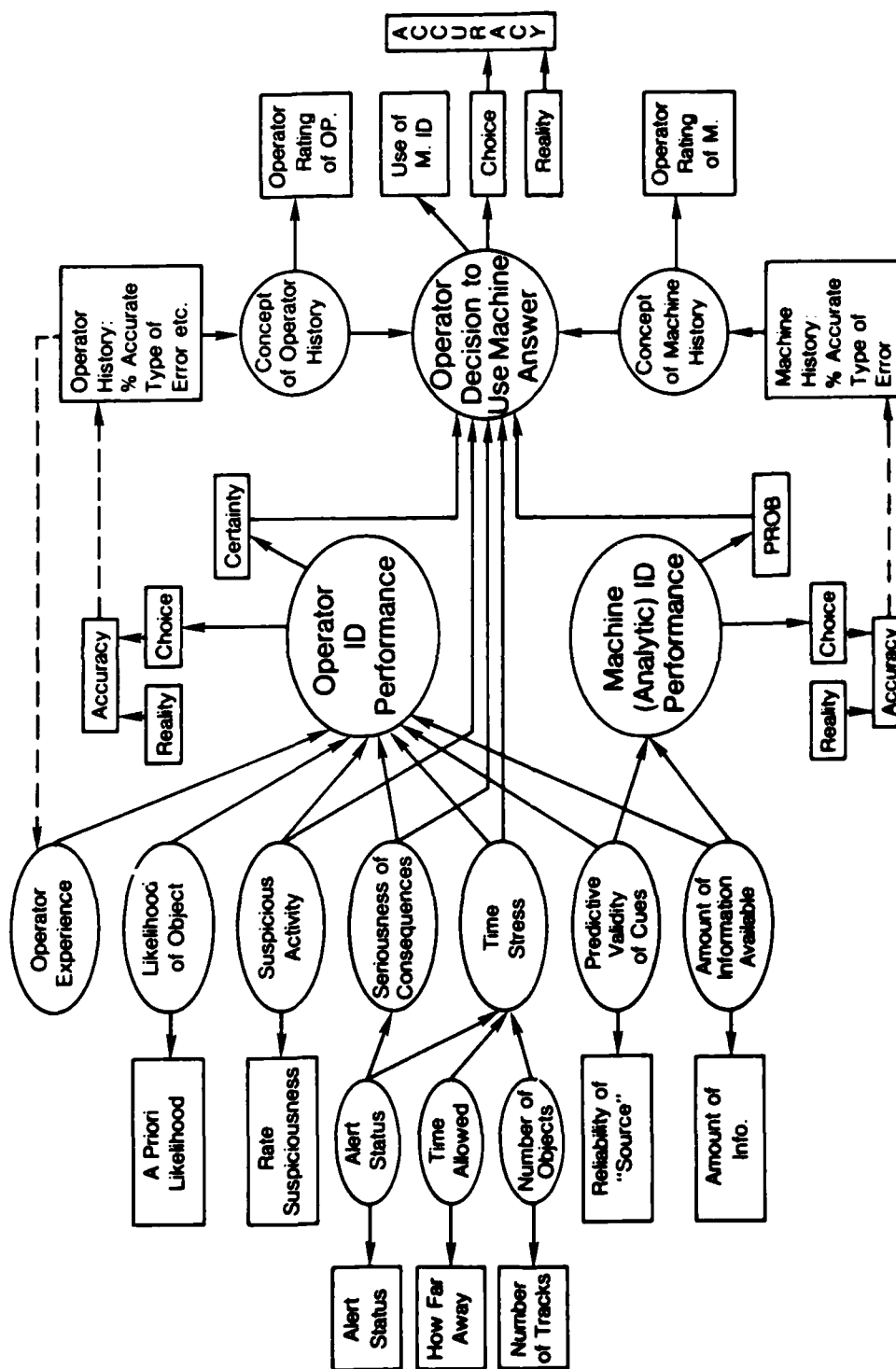


Figure 5. Implementation of Automation in Inference Task

## VI. RECOMMENDATIONS:

The previous section outlined hypothesized effects of situational constraints upon inference process and performance. In addition, a preliminary predictive model was constructed relating the implementation of an automated subsystem within a "manual" inference system. The relationship between the variables need to be empirically obtained to confirm the theoretical assumptions.

I propose to do this by developing a laboratory inference task that has the characteristics outlined earlier (complex, sequential cue acquisition, etc.). This inference task will be run on a computer terminal so that the subject can request cues and respond with hypotheses and certainty ratings. This will allow collection of data relevant to the process and performance variables outlined previously.

For assessment of the processes used in a manually performed inference task, situational constraints will be manipulated and their effects on process and performance variables measured. This will allow determination of what processes are utilized under various task conditions.

A second phase of the follow-on research will be to provide the subject with a computer aid. This will be a simulated computer aiding system because the answer to be given to the subject will be predetermined by the Experimenter. In this way, the accuracy of the answers provided to the perceiver can be manipulated as an independent variable. It is expected that as accuracy of the automated system increases, the use of the machine will also increase in an exponentially increasing fashion. Situational constraints will also be varied similarly to the design used in assessment of the completely manual inference process.

# REFERENCES

1. P. Slovic, B. Fischhoff, and S. Lichtenstein, "Behavioral Decision Theory," Annual Review of Psychology, Vol. 28, pp. 1-39, 1977.
2. H.E. Price, R.E. Maisana, and H.P. Van Cott, "The Allocation of Functions in Man-Machine Systems: A Perspective and Literature Review," (NUREG CR-2623). Oak Ridge, Te: Oak Ridge National Laboratory, June 1982.
3. D.B. Yntema and W.S. Torgerson, "Man-Computer Cooperation in Decisions Requiring Common Sense," IRE Transactions on Human Factors in Electronics, Vol. 2, pp. 20-26, 1961.
4. H.J. Einhorn and R.M. Hogarth, "Behavioral Decision Theory: Processes of Judgment and Choice," Annual Review of Psychology, Vol. 32, pp. 53-88, 1981.
5. P. Slovic and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," Organizational Behavior and Human Performance, Vol. 6, pp. 649-744, 1971.
6. H.J. Einhorn, D.N. Kleinmuntz, and B. Kleinmuntz, "Linear Regression and Process-tracing Models of Judgment," Psychological Review, Vol. 86, pp. 465-485, 1979.
7. M.W. Schustack and R.J. Sternberg, "Evaluation of Evidence in Causal Inference," Journal of Experimental Psychology: General, Vol. 110, pp. 101-120, 1981.
8. E.B. Ebbeson and V.J. Konecni, "On the External Validity of Decision-Making Research: What Do We Know About Decisions in the Real World?," In T.S. Wallsten (Ed.), Cognitive Processes in Choice and Decision Behavior, Hillsdale, N.J.: Lawrence Erlbaum, 1980.
9. H.J. Einhorn, "Expert Judgment: Some Necessary Conditions and An Example," Journal of Applied Psychology, Vol. 5, pp. 562-571, 1974.
10. R.H. Ashton, "Cue Utilization and Expert Judgments: A Comparison of Independent Auditors with Other Judges," Journal of Applied Psychology, Vol. 59, pp. 437-444, 1974.
11. D.L. Medin, M.W. Altom, S.M. Edelson, and D. Freko, "Correlated Symptoms and Simulated Medical Classification," Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 8, pp. 37-50, 1982.
12. A.S. Elstein, L.E. Shulman, and S.A. Sprafka, Medical Problem Solving: An Analysis of Clinical Reasoning. Cambridge, Mass: Harvard Univ. Press, 1978.
13. J.W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance, Vol. 16, pp. 366-387, 1976.



14. D.M. Green and J.A. Swets, Signal Detection Theory and Psychophysics. New York: Wiley, 1966.
15. J.A. Swets and R.M. Pickett, Evaluation of Diagnostic Systems: Methods from Signal Detection Theory. Cambridge, Mass: Bolt, Beranek and Newman Inc., 1982.
16. S.G. Hart and T.B. Sheridan, "Pilot Workload, Performance, and Aircraft Control Automation," Proceedings of the AGARD Symposium on Human Factors Considerations in High Performance Aircraft, Williamsburg, VA; April, 1984.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ANGLE RESOLVED ION-SCATTERING SPECTROSCOPY-A FEASIBILITY STUDY

Prepared by:	Thomas P. Graham, Ph. D.
Academic Rank:	Professor
Department and University:	Physics Department, University of Dayton
Research Location:	Wright Patterson Air Force Materials Laboratory, Mechanics and Surface Interactions Branch, Surface Group
USAF Research:	William L. Baun
Date:	September 17, 1984
Contract No:	F49620-82-C-0035

ANGLE RESOLVED ION-SCATTERING SPECTROSCOPY-

A FEASIBILITY STUDY

by

Thomas P. Graham

ABSTRACT

An angle resolved ion-scattering system was assembled, tested and calibrated. The system performed well in both the constant transmission mode and the sector sweeping mode. Samples of gold, GaAs coated with gold and GaAs were studied using helium, neon and sodium ion beams accelerated to a number of different voltages. Angular studies were carried out over a limited range of angles. Good agreement was obtained with available theory in most cases. A number of deficiencies of the system were discovered and suggestions are made to remedy them.

### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a meaningful and interesting summer at the Air Force Materials Laboratory at Wright Patterson Air Force Base in Dayton, Ohio. In particular , he would like to acknowledge the Mechanics and Surface Interactions branch and T. W. Haas for their hospitality.

Finally, he would like to to thank William Baun for suggesting this area of research and for his many helpful suggestions and collaboration through the summer. Thanks go also to R. Twist, J. Solomon, L. Grazulius and J. Grant.

## I. Introduction:

The characterization of the surface properties of electronic materials is a very important requirement for the successful design of semiconducting devices for a number of Air Force programs. At present there is a particular interest in GaAs for use in detectors and FET devices. The characterization includes quantitative and qualitative determination of surface species as well as surface structure.

Ion scattering spectroscopy (ISS), the scattering of low energy ions in the range of 0.2 to 10 kev, exhibits excellent surface selectivity. The ion scattering that results is attributed to scattering from the first one or two surface layers of the sample. This selectivity arises from: (a) large scattering cross-sections which leads to depletion of the ion beam as it enters and leaves the solid, and (b) to a neutralization effect whereby ions penetrating beyond the first layer of atoms at a solid surface are neutralized very efficiently.<sup>1</sup>

When a beam of accelerated ions strikes a crystal surface, each encounter between a projectile ion and a surface atom can be described in terms of an elastic binary collision. Under certain circumstances multiple scattering, blocking, shadowing and channeling effects become important. The energy of an ion making an elastic collision, expressed as a ratio of final energy,  $E_1$  to initial energy,  $E_0$  is determined from energy and momentum conservation to be:

$$E_1/E_0 = \left[ M_1^2 / (M_1 + M_2)^2 \right] \left[ \cos^2 \theta_L + (M_2^2 / M_1^2 - \sin^2 \theta_L)^{1/2} \right]^2 \quad (1)$$

Here  $M_1$  is the mass of the incoming ion,  $M_2$  the mass of the target atom and  $\theta_L$  the laboratory scattering angle. By measuring the energy distribution of the scattered ions, the elements in the region of impact can be determined.<sup>1</sup>

While ISS is used routinely for elemental characterization, interest has grown in recent years in using angle resolved ion scattering (ARISS as it is called by some groups<sup>2</sup>) to determine surface structure often in conjunction with LEED measurements. Analysis of

multiple scattering effects and of shadowing and blocking effects can be used to construct models of the surface. When multiple collisions occur, shoulders or peaks occur at energies greater than that of the binary peak for a given element. Since different collision sequences lead to different final energies, knowledge of the energies and angular distributions in principle enable structural analysis. When shadowing, blocking, focussing or channeling effects occur, the intensities of the binary peaks will depend on the energies and incident and final angles.

Heiland and Taglauer<sup>3</sup> studied shadowing effects of adsorbed species on metal surfaces to estimate the positions of the adsorbed atom on the surface. Algra et al<sup>4</sup> determined the structure of a stepped Cu surface using multiple collision analysis. Overbury et al have investigated the structure of Au(110)<sup>2</sup> and adsorbate ordering on Mo(001)<sup>5</sup> using ARISS. Bronckers and de Wit have studied the first two layers of Cu(110) using shadowing effects.<sup>6,7</sup> Marchut et al<sup>8</sup> have analysed an unreconstructed Fe(001) surface using a shadow cone analysis involving the second and third atomic layers.

## II. Objectives

The main objective of the project was to assemble, test and gain experience with a rudimentary angle resolved ion scattering spectrometer. The experience with this system, particularly with its limitations and problems, will provide the basis for the design of a more versatile and sophisticated ARISS system.

A secondary objective was to compare the results obtained with rare gas ions and alkali-metal ions. Neon and sodium were to be scattered off gold and gallium arsenide surfaces in this phase of the work.

## III. DESCRIPTION OF THE ARISS SYSTEM

A schematic for the angle resolved ion scattering spectrometer (ARISS) system is shown in Figure 1. The heart of the system is its dispersive element, the electrostatic analyzer. The analyzer provides a

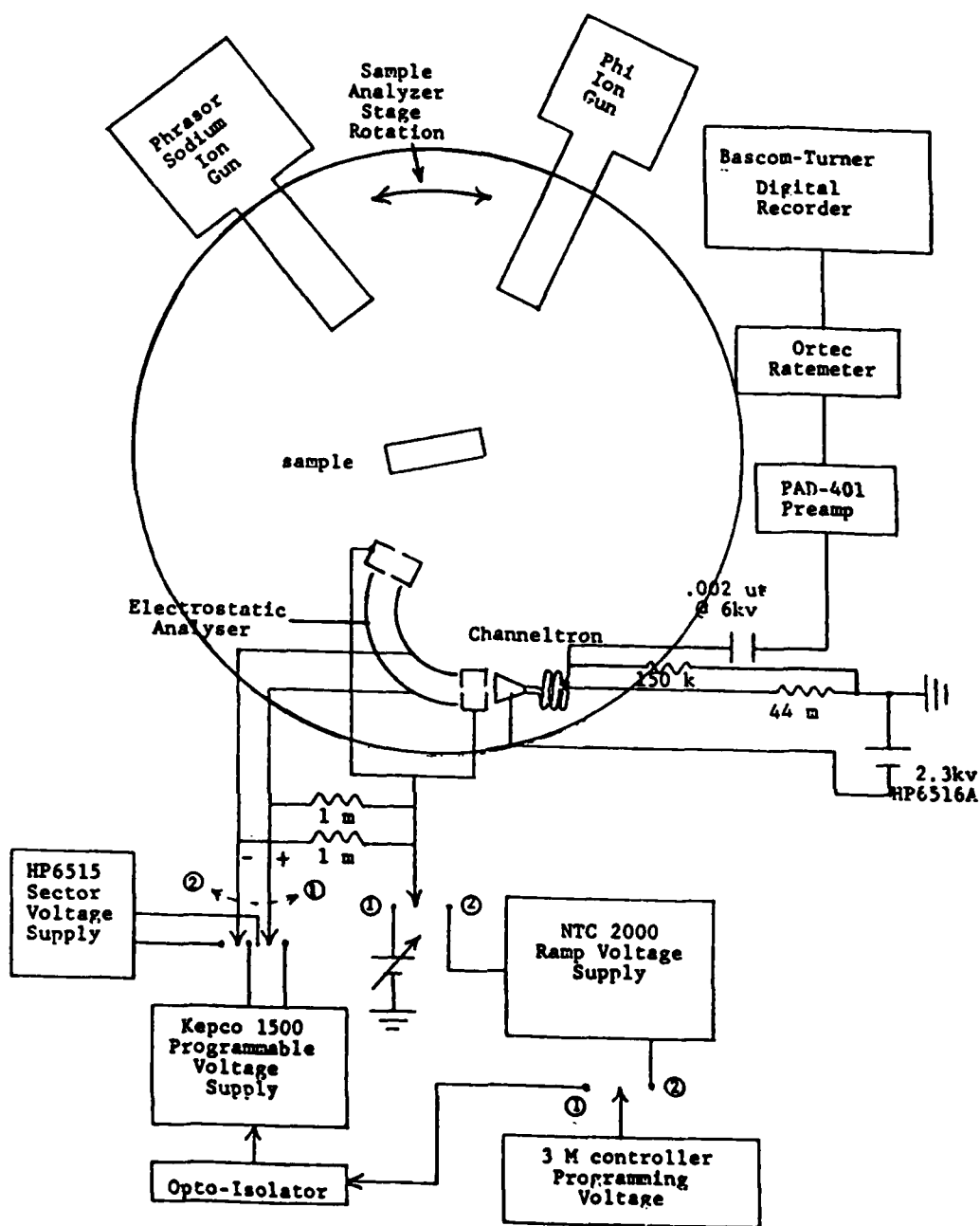


Figure 1. ARISS system schematic

means for analyzing the energy distributions of ions incident upon its entrance aperture. Particles within an energy range will be transmitted between the inner and outer spherical sector surfaces of the analyzer and refocused upon the exit aperture. By sweeping the reference potential of the whole analyzer with a fixed potential across the hemispheres, or by sweeping the potential across the hemispheres with a fixed reference potential, the input energy spectrum is obtained.

The analyzer is a Comstock Double Focusing Electrostatic Analyzer Model AC-901.<sup>9</sup> It consists of two concentric 160 degree spherical sector surfaces; an inner convex one of radius 3.25 cm, and an outer concave surface of radius 4.05 cm. Two end plates are provided with knurled screw caps for easy insertion and changing of the apertures. The focal points of the analyzer lie at the entrance and exit apertures. Single charged particles of energy  $E_p$  are transmitted between concentric constant potential spheres of radii  $r_1$  and  $r_2$  if the potential difference,  $V_H$ , between the spheres satisfies

$$E_p = V_H / (r_1/r_2 - r_2/r_1) \quad (2)$$

$$\text{In this case} \quad E_p = V_H / C \quad \text{where } C = 1/2.254 \quad (3)$$

The analyzer can be operated in either of two modes as indicated above. These are (1) the sector sweeping mode and (2) the constant transmission (resolution) mode. The electrical connections and power supplies used for each mode are shown in Figure 1.

In the sector sweeping mode, equal and opposite voltages are applied to the sector surfaces and the apertures and side plates are held at a fixed voltage, often ground potential. As the sector voltage,  $V_H$ , is swept, the energy distribution is obtained. In this mode transmission efficiency and absolute energy resolution change as  $V_H$  is swept. It is a useful mode when external focusing conditions are sensitive to the potential on the apertures.

In the constant transmission mode, the sector voltage,  $V_H$ , is fixed while the voltage on the apertures and side walls are swept. The ions are either accelerated or retarded until their energy matches the



transmission energy at which they are focused on the detector. In this mode both transmission and resolution are constant for a given  $V_H$ . By choosing  $V_H$  small, high resolution (but lowered count rate) can be obtained. The manufacturer gives the energy resolution for 1 mm diameter apertures as 0.8% and for 0.5 mm diameter apertures as 0.4%.

Once the ions traverse the analyzer they are detected by a Channeltron detector (model CEM 4013). The cone of the Channeltron is held at about 2300 volts negative so the efficiency for electron multiplication is about the same for all ion energies used. The current pulses are coupled via a capacitor to a pre-amp and the signal acquisition equipment.

The sample and the analyzer are rigidly attached to a circular plate which can be rotated relative to fixed ion guns. An angular scale is scribed on the plate. The axis of rotation is co-linear with the center line of the long side of a tee with a six inch inner diameter which forms the experimental chamber. One end of the chamber accommodates a base plate with a rotary feedthrough to control the rotation of the analyzer-sample stage, and the electrical feedthroughs. The other end has a window. Two ports, 90 degrees apart, with axes perpendicular to the rotation axes hold the ion guns, a sodium gun and an ion gun of the electron impact type. Helium and neon gases were used with this gun. A four inch window was on a port half-way between the ion gun ports to enable angle measurements to be made with the aid of a reference pin. This chamber was appended to a larger chamber which could be pumped with sorb pumps, an ion pump and a sublimation pump. The base pressure for the system was about  $2 \times 10^{-8}$  torr. When helium or neon was used the system was backfilled to  $2 \times 10^{-5}$  torr.

An example of a spectrum taken with the system is shown in Figure 2. It is the spectrum of sodium scattered from a GaAs sample with a thin coating of gold. The scattering angle was 80 degrees and the initial ion energy was 2 keV. There is apparent multiple scattering at energies above the binary collision peaks and a rather high intensity below the peaks due to inelastic losses which result from scattering from layers below the surface. This latter effect occurs more for alkali atom scattering because of the much smaller neutralization probability.

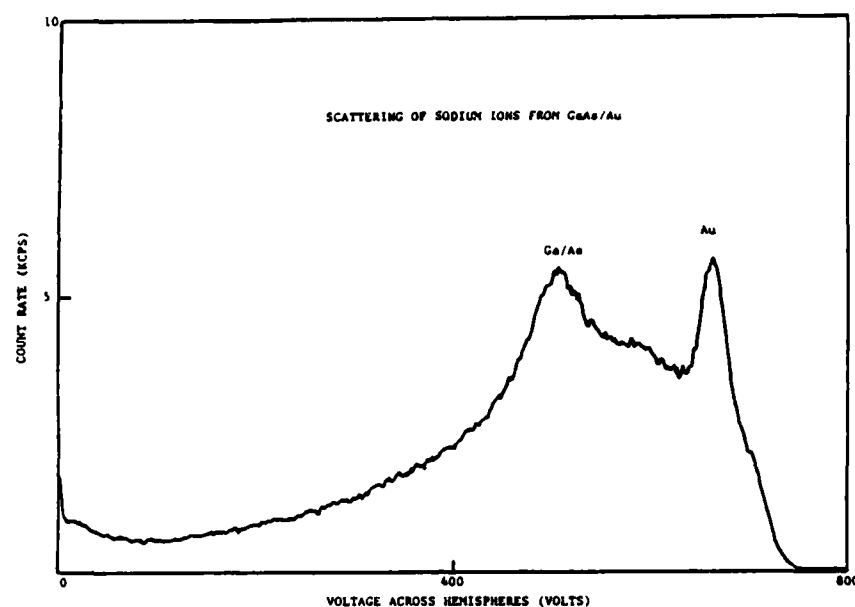


Figure 2. Spectrum of Na scattered from GaAs/Au

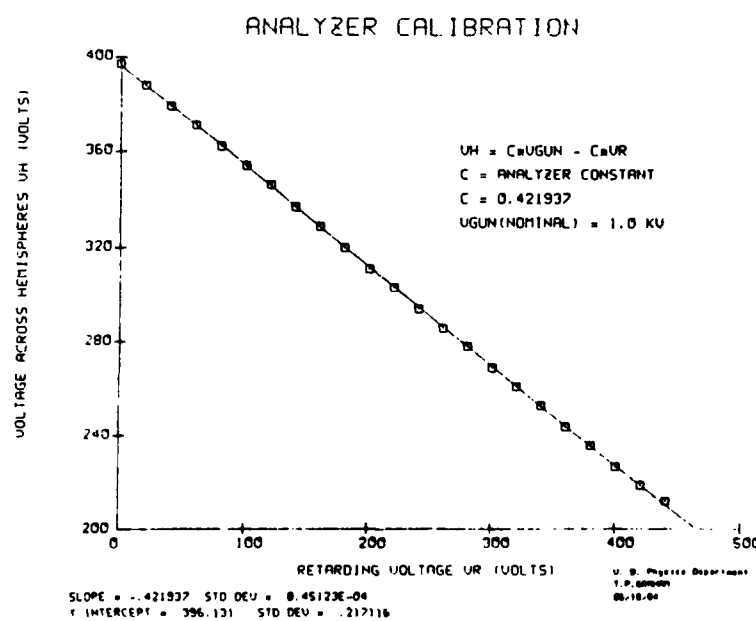


Figure 3. Analyzer calibration

#### IV. SYSTEM PERFORMANCE

The ARISS system was tested extensively using a variety of experimental conditions. Both modes were tested, using helium, neon and sodium as the projectile ions. The test encompassed a range of accelerating voltages from about 500 V to 2500 V. The tests used either no sample, i.e. the ions fired directly into the analyzer, an aluminum block coated with gold, a GaAs sample coated with gold or a GaAs (110) surface. In this section, a selection of the results will be presented to illustrate the behavior of the system.

The energy of an ion which approaches the analyzer and is subsequently detected,  $E$ , is just the sum of the pass energy,  $E_p$ , and the energy gained or lost as a result of the accelerating or retarding voltage which will be denoted by  $V_R$ .  $E_p$  is given by Equation 3. If the incoming ion has been scattered off a surface in a binary collision,  $E$  is just the  $E_1$  of Equation 1. In general, we could write  $E = B \cdot E_0$  where  $B$  would depend on the details of the scattering process involved and is a function of angle. Since  $E_0$  results from an acceleration in the ion gun, it is convenient to write it as  $V_{GUN}$  (in eV's). With these definitions we can then take as a working equation for the analyzer:

$$E = B \cdot V_{GUN} = V_R + V_H/C \quad (4)$$

With this equation we can analyze experiments using either mode for the analyzer constant,  $C$ , or the energy of the incident ion,  $V_{GUN}$  or the energy ratio,  $B = E/E_0$ .  $B = 1$  corresponds to the case of no scattering.  $V_R$  equal to zero or a constant describes the sector sweep mode and  $V_H$  constant, the constant transmission mode.

Another useful equation is based on the definition of the resolution,  $R$ , of the analyzer,  $\Delta E/E$ . If  $\Delta E_0$  is the linewidth (FWHM) of the beam incident on the analyzer, then the detected linewidth including the contribution from the analyser energy distribution can be written as:

$$\Delta E = \Delta E_0 + R \cdot E \quad (5)$$

The analyzer constant, C, was initially determined by using neon gas and the Phi ion gun and detecting the ion current to the cone of the detector. The voltage across the sectors was set and the retarding voltage varied until the ions traversed the analyzer and exited to the detector. A plot of the data is given in Figure 3. The solid line is a least squares fit and gives 0.422 for the constant which is about 5% lower than given by the manufacturer. However the analyzer had been disassembled a number of times and the spacing and alignment had probably changed. As the retarding increased, the beam had to be refocused, indicating the need for a better shielding of the beam region from the potentials. This was subsequently done but the problem was not completely eliminated.

Figures 4 and 5 show the results obtained by firing neon and sodium ions directly into the entrance aperture of the analyzer. The analyzer constant was obtained in each case. In the neon case the average constant was about 0.3% lower than the value given above while the sodium data gave a value about 1.3% lower. At zero and low retarding voltages, the width of the transmission peak was quite small. But as higher retarding voltages were used, the line shape became distorted and even multiple peaks were observed. This seems to imply a focusing problem caused by the potentials on the entrance aperture. This data also illustrates the way the initial ion energy can be determined from the intercept on the ordinate.

From the line width obtained with neon fired directly at the analyzer with 0.5 mm diameter apertures at an energy of 1.5 keV with zero retardation, the resolution was found to be 0.63%. This compares favorably with that quoted by the manufacturer. Because of the distortion of the line shape with retarding voltage a least squares analysis of the retarding data was not possible.

Figure 6 is a plot of the line widths obtained from neon scattered off gold using 1 mm diameter apertures. The slope is, from Equation 5, R/C. A value for R of 1.3% is obtained which is comparable to that given by the manufacturer. From the same experiment, the count rate as a

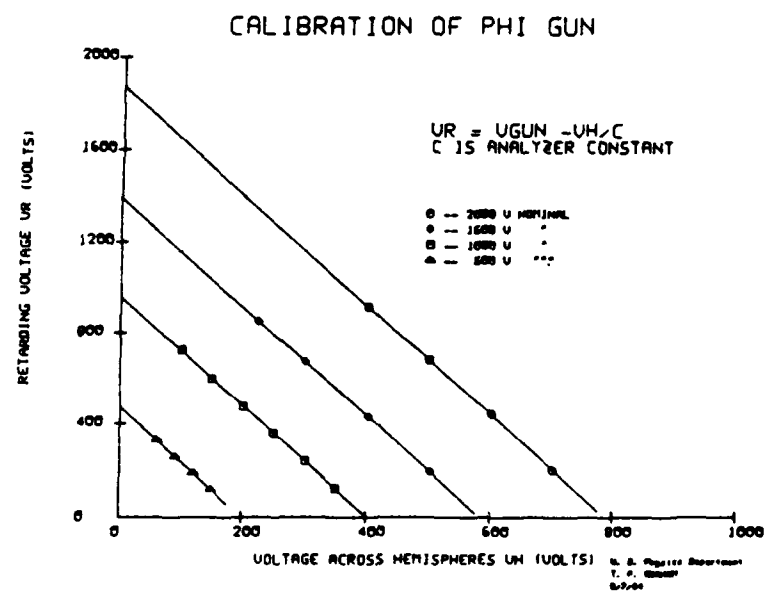


Figure 4. Calibration of Phi gun

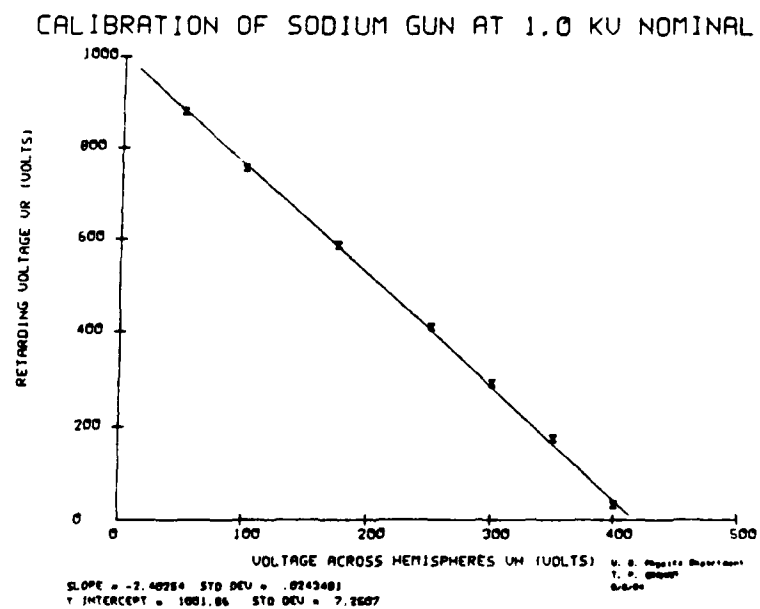


Figure 5. Calibration of sodium gun

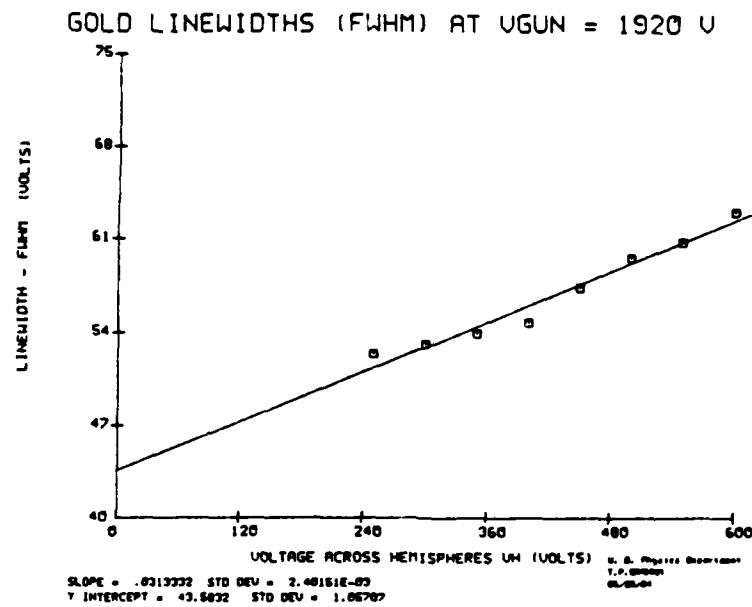


Figure 6. Linewidths of gold peak from neon scattering

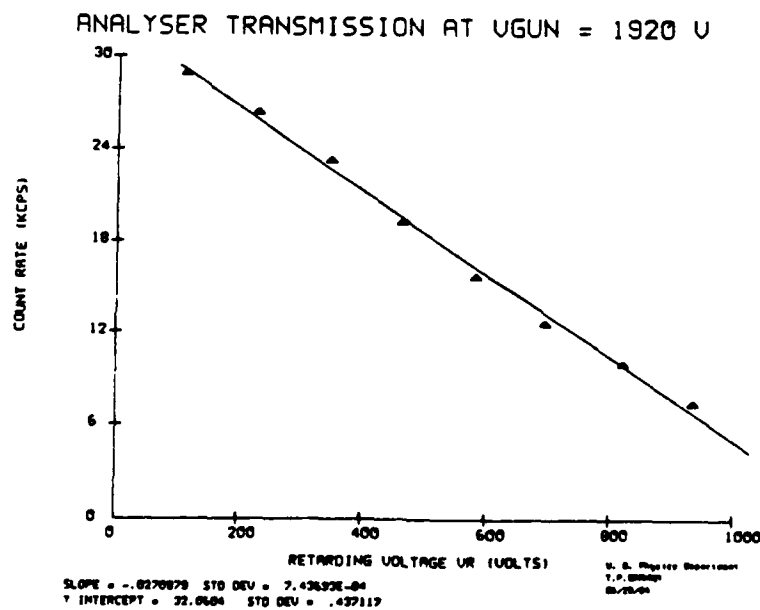


Figure 7. Analyzer transmission

function of retarding voltage can be obtained. This is shown in Figure 7. The linear fall-off is in accord with the predictions of Helmer.<sup>10</sup> The higher retarding voltages correspond to smaller voltages across the sectors. This implies that in the constant transmission mode, for high resolution (small  $V_H$ ) we would obtain lower count rates.

Figures 8 and 9 illustrate how the scattered energy and analyzer constant can be extracted from the data for scattering from a sample. The constant is essentially the same as obtained by directly aiming the gun into the analyzer. The energy ratios are about 3% lower than expected on the basis of Equation 1 for the angle used ( $85^\circ$ ).

In sum, the system performs substantially as expected with some limitations as a high resolution instrument.

#### V. ANGULAR STUDIES

The angular dependence of the scattering of both neon and sodium from a number of samples was investigated. Because of the compactness of the system, data could be taken for only about a range of 65 degrees before the shielding enclosure of the analyzer interfered with the ion beam. Also since the sample could not be moved independently of the analyzer, the angle of reflection and the azimuthal angle could not be changed. It was possible to obtain data to compare with Equation 1 and to see as yet unexplained intensity variations as the scattering angle was varied.

Figure 10 shows data taken in the constant transmission mode ( $V_H=150$  V) for neon on gold. 35 degrees should be subtracted from the angles given to get the scattering angle. The X's go with the right hand scale and show the intensity variation with angle. The retarding voltage data, boxes and left scale, is consistent with Equation 1. How much of the intensity variation is due to the changing retarding voltage is not known. Some of the variation at the smaller angles is undoubtedly due to the fact that more of the ion beam impinges on regions of the sample not viewed by the entrance aperture.

# SCATTERING OF NEON FROM GOLD AT UGUN = 1442 V

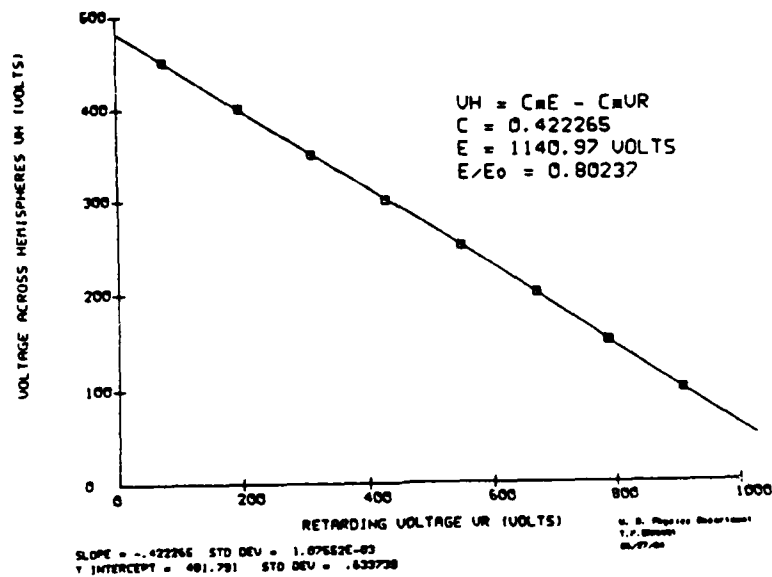


Figure 8. Scattering of neon from gold

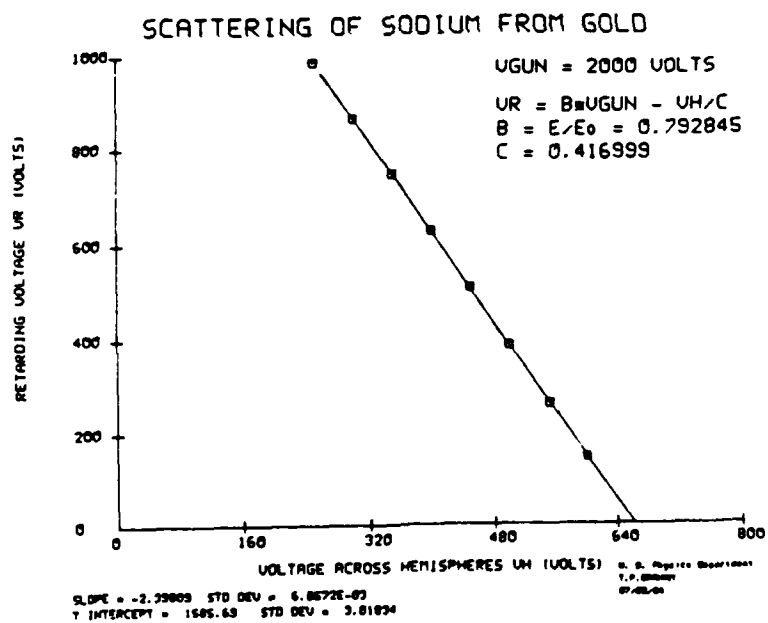


Figure 9. Scattering of sodium from gold



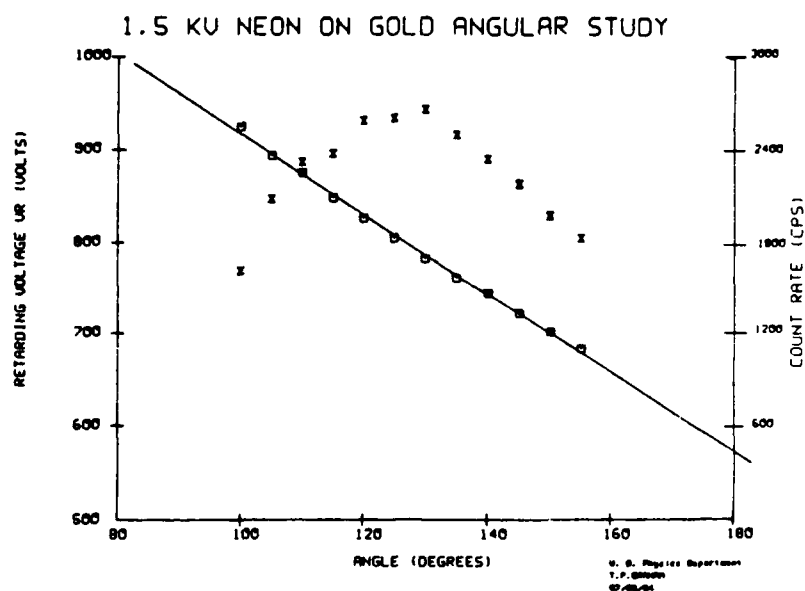


Figure 10. Angular variation of retarding voltage and intensity

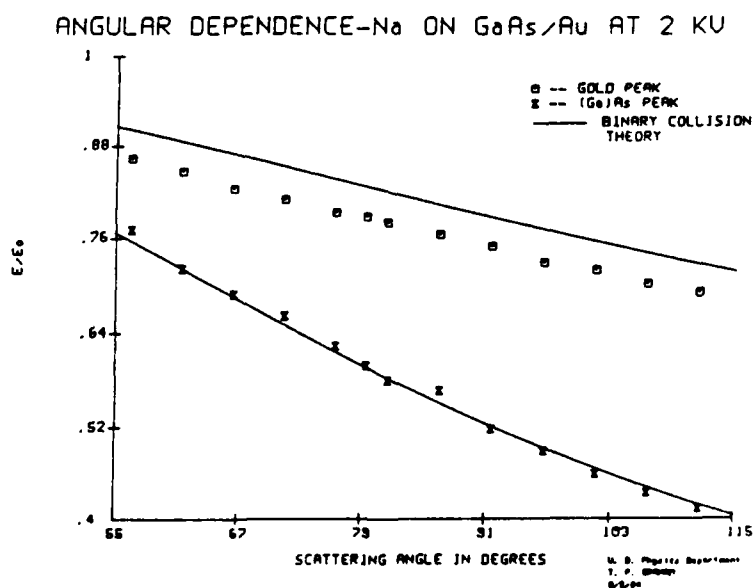


Figure 11. Angular dependence of Na scattered of GaAs/Au

Figure 11 shows data obtained from the scattering of sodium off a GaAs sample with a coating of gold. The Ga and As peaks were not separately resolved as seen in Figure 2. Comparison is made with the ratios calculated on the basis of Equation 1. The data for gold is consistently low. The gold peak is sitting on the multiple scattering shoulder of GaAs which would tend to lower the apparent peak position. Also recent studies on silver found the peak to be 1 to 2% low due to an inelastic energy loss mechanism.<sup>11</sup>

The angular dependence of the energy ratio for scattering of neon from a GaAs surface is shown in Figure 12. Good agreement is seen with the binary collision model. From the same experiment, the intensity vs angle dependence was determined and is shown in Figure 13. The solid lines represent third order polynomials fit to the data to help guide the eye. The count rates were rather low and there is noticeable scatter to the data.

## VI. RECOMMENDATIONS

Within the limitations of time and available components and equipment the project was successful. From it we can point to a number of improvements which must be made in order to develop a high quality ARISS system.

1. The system should be built in a large vacuum chamber- 12 to 15 inches in diameter. This will allow the analyzer to be placed a relatively large distance from the scattering region. A much larger range of scattering angles would be allowed and extraneous fields in the scattering region would be reduced. Room for other instruments such as LEED would be available.
2. An input lens system should be added to the system to collect the ions and focus them on the entrance aperture. This would allow operation with a relatively large aperture for sensitivity while maintaining high resolution. An appropriate lens would permit operation in the constant transmission mode with the sector voltage low enough to obtain high resolution.

# ANGULAR DEPENDENCE - NEON ON GaAs AT 1.5 KV

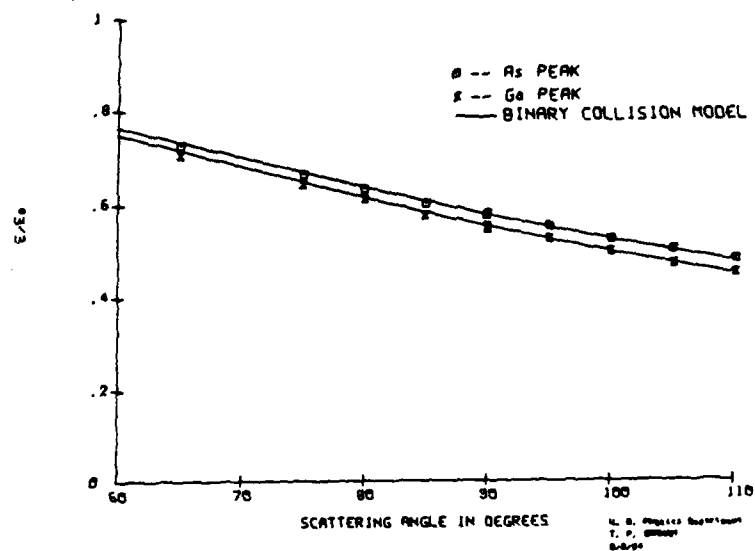


Figure 12. Angular dependence of neon scattered from GaAs

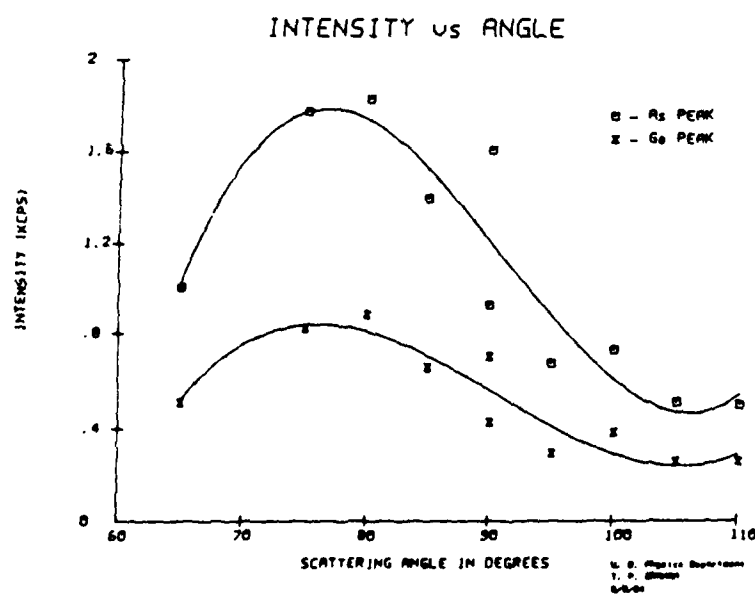


Figure 13. Angular dependence of intensity for neon scattered from GaAs

An output lens, while less necessary, would be useful in increasing detection efficiency and reducing noise by the use of a smaller detector aperture.

3. A sample manipulator to allow placement of the sample in any orientation is needed to be able to map out the intensity variations in various crystallographic directions.
4. Complete shielding of the entire analyzer-channeltron assembly, which would be possible in a larger chamber, is necessary.

Acting on the above recommendations could form the basis for follow-on research. If, in addition the capability to do laser multiphoton resonance ionization were added, the sensitivity of the experiments could be enhanced by preferentially detecting certain of the sputtered ions.

### References

1. Buck, T. M., in Methods of Surface Analysis, Ed. by A. W. Czanderna, (Elsevier, Amsterdam, 1975), Ch. 3.
2. Overbury, S. H., Heiland, W., Zehner, D. M., Datz, S. and Thoe, R. S., Surface Sci., Vol. 109, pp. 239-262, 1981.
3. Heiland, W. and Taglauer, E., Surface Sci., Vol. 68, pp. 96-107, 1977.
4. Algra A. J., Luitjens, S. B., Suurmeijer, E. P. Th. M. and Boers, A. L., Surface Sci., Vol. 100, pp. 329-341, 1980.
5. DeKoven, B. M. and Overbury, S. H., Phys. Rev. Lett., Vol 53, pp.481-484, 1984.
6. Bronckers, R. P. N. and de Wit, A. G. J., Surface Sci., Vol. 104, pp. 384-404, 1981.
7. Bronckers, R. P. N. and de Wit, A. G. J., Surface Sci., Vol. 112, pp.111-132, 1981.
8. Marchut, L., Buck, T. M., Wheatley, G. H. and McMahon, C. J. Jr., Surface Sci., Vol. 141, pp.549-566, 1984.
9. Comstock, Inc. P. O. Box 199, Oak Ridge, Tenn. 37830
10. Helmer, J. C., Am. J. Phys., Vol. 34, pp.222-227, 1966.
11. MacDonald, R. J. and O'Conner, D. J., Surface Sci., Vol. 124, pp. 433-439, 1983

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ELECTROGASTROGRAM AND ITS EFFECTIVENESS  
IN EVALUATION OF MOTION SICKNESS

Prepared by:	Edward Carl Greco, Jr., Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Biomedical Engineering University of Miami
Research Location:	Neuropsychiatry Clinical Sciences Division School of Aerospace Medicine
USAF Research:	David R. Jones, Col, USAF, MC, CFS
Date:	10 August 1984
Contract No:	F49620-82-C-0035

ELECTROGASTROGRAM AND ITS EFFECTIVENESS  
IN EVALUATION OF MOTION SICKNESS

by

Edward Carl Greco, Jr.

ABSTRACT

The electrogastrogram (EGG) is a low frequency signal recorded from the abdomen with surface electrodes which is thought to represent gastric motility. Its frequency content ranges from 1 to 5 cycles/min. The EGG offers the potential for a direct observation of gastric motility which would be very useful as an index of nausea during motion stimuli. The EGG could be used as a biofeedback variable providing an indicator of gastric activity. Subjects monitoring their own EGG may be able to voluntarily control gastric activity during motion stimuli and thus learn to minimize their susceptibility to motion sickness. However, there are several problems associated with the recording techniques and interpretation of results which requires further investigation before the EGG can be used effectively for biofeedback treatment of motion sickness.

### Acknowledgements

I wish to thank the Air Force System Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for arranging a very interesting and enlightening summer at the School of Aerospace Medicine, Brooks AFB, Texas. In particular I would like to thank the personnel in the Neuropsychiatry Branch of the Clinical Sciences Division for their assistance and hospitality.

I offer a special thanks to Dr. David R. Jones for providing the research project, guiding, and supporting my research efforts; and to Dr. Bryce Hartman for the "big picture" orientation and his administrative assistance. Thanks also to Lt. James Heriot for his assistance in the first half of this project and to Mr. John Davis for his technical assistance throughout the summer. Thank you Drs. Giles, Lewis, Patterson, and Perrien for your helpful discussions and suggestions during my summer tenure. Finally, thanks to the technical staff of Neuropsychiatry for your assistance whenever asked and for your friendliness which made the summer a very enjoyable experience.



## I. Introduction.

The debilitating effect of motion/air sickness may adversely and severely affect fliers ability to function. Student, as well as experienced aviators, are susceptible to the malaise of this unpredictable affliction [Levy, Jones, Carlson, 1981]. A form of motion sickness has also been experienced by approximately 40% of all astronauts. Astronauts were more susceptible to space adaption syndrome (SAS, as it is now called) the first several days in space but symptoms persisted for up to two weeks [Oman, 1982]. The seriousness of motion sickness for space, as well as air, operations has been recently publicized in the popular press [Chaikin, 1984; Joyce, 1984].

Symptoms of motion sickness include hypersalivation, facial pallor, sweating, warm flush, malaise, headaches, drowsiness, epigastric awareness, epigastric discomfort, nausea, and vomiting. Epigastric awareness is defined as the feeling which draws attention to the epigastric area but is not uncomfortable, while the feeling of distress which is more than awareness but less than nausea has been termed epigastric discomfort [Graybiel, 1968]. In addition there are symptoms which have not been regularly reported. These include: (1) symptoms which are difficult to measure, i.e., respiratory irregularities and electrical activity of the

stomach and intestine; and (2) symptoms which have not been shown to be reliable, e.g., blurred vision [Graybiel, 1968].

The treatment for motion sickness has evolved around two paradigms: pharmacological and self-asserted autonomic control. Pharmacological agents include drugs which act on the cholinergic receptors. Autonomic control has been taught through biofeedback techniques. Scopolamine and scopolamine combined with amphetamine are anti-cholinergic drugs which have been used to treat motion and space sickness. These drugs are short acting but are ineffective if administered after symptoms have appeared. The short term side effects of scopolamine include central nervous system depression, dry mouth, and loss of visual accommodation as well as longer term effects such as memory impairment, sleep disturbances, and prenatal toxicity [Kohl, 1983]. Dexedrine has been combined with scopolamine and amphetamine in order to counteract the drowsiness side effect. However these drugs can not be used by solo pilots because of the potential side effects [Jones, 1984]. Some of the side effects may be reduced with alternate administration procedures; for example, a skin patch for absorption across the skin. Skin absorption administration reduces the peak intravenous concentration and prolongs the drugs effectiveness. However these drugs do not totally eliminate the motion sickness but simply raise the crew person's tolerance to motion stimuli

[Oman, 1982].

Volitional control over autonomic functions during motion stimuli has been achieved with biofeedback conditioning [Cowings, 1977]. Biofeedback involves the measurement and presentation of physiological variables to a subject. Physiological signals which have been found to be sensitive to subjects' autonomic state during motion stimulus and can be readily measured are the following: (1) skin temperature [Levy, 1981], (2) galvanic skin response [Levy, 1981], (3) frontalis EMG [Levy, 1981], (4) heart rate [Cowings, 1977], (5) respiration rate [Cowings, 1977], and (6) blood volume pulse of the face and hands [Cowings, 1977]. Subjects are able to gain control over their autonomic state by voluntarily shifting these recorded physiological variables to some desired range and thus reduce their sensitivity to motion stimuli [Cowings, 1977; Levy, 1981; Jones, 1984].

## II. Objectives.

The objectives of this research study were to evaluate existing instrumentation and equipment; and to specify and design additional instrumentation to monitor motion stimuli responses in order to provide biofeedback training to increase motion stimulus tolerance. Existing laboratory equip-

ment and instrumentation in the motion stimulus laboratory of Neuropsychiatry, and instrumentation provided by the Air Force Institute of Technology [Earl, 1983], under separate contract, were evaluated, modified, and integrated into a laboratory system for motion sickness evaluation and biofeedback training. This report concentrates on the electrogastrogram (EGG): a physiological variable which has not received widespread interest for evaluation of motion stimuli until recently [Earl, 1983; Patterson, 1967].

### III. Electrogastrogram.

Low frequency electrical activity recorded from skin surface electrodes on the abdomen has been designated the electrogastrogram (EGG). The EGG was thought to originate from the smooth musculature of the lower alimentary canal, including the stomach and intestine [Russell, 1967]. Electrical activity of the gastrointestinal tract has been recorded from electrodes places directly on exposed or excised tissue in several early investigations. A close correspondance between the electrical and mechanical activity in both amplitude as well as form was observed in these studies [Russell, 1967]. An intragastric electrode technique was developed by Goodman [1942] with an electrode placed into the stomach through a nasal gastric tube and a refer-

ence electrode placed on the forearm. This intragastric electrode technique formed the basis for several clinical studies in which electrical activity was found to correlate with intragastric balloon pressure measurements [Russell, 1967]. The first attempt to record g.i. activity from external surface electrodes on humans was made by Alvarez [1922]. These early human EGG's were not very successful mainly due to the lack of low-frequency, high-gain amplifiers [Russell, 1967].

The EGG was recorded with silver disk electrodes which had been coated with silver/silver chloride to minimize polarization effects. The subject's skin was prepared to accept the electrodes by shaving and either rubbing the shaved area with electrode jelly or lightly abrading with fine sandpaper. The electrodes were coated with electrode jelly prior to attaching to the subject [Russell, 1967]. The reference electrode was placed on the subject's leg with the active electrode placed in one of four quadrants sites on the abdomen. The standard site for the active electrode was 1 in. above and 2 in. to the left of the navel. Another site was 1 in. above and 2 in. to the right with other sites symmetrically located around the navel.

The EGG recording was observed to depend on several factors including: electrode placement, content of stomach,

body position, drugs, visual and auditory stimulation, smoking, and task required of subject [Russell, 1967]. There were common features among all recordings which could be classified into three types. Type I signals were low magnitude, single frequency waves with frequencies in the range of 1 to 2 per min in the fundus of the stomach and 3 to 4 per min in the antrum. Type II signals were larger amplitude waves than type I. The harmonic content of type II was simple, like type I, but the frequency was time varying. The frequency covered the range from 2 to 5 per min during non-rhythmic activity and 3 per min during rhythmic conditions. Type III waves were complex with frequency components between 1 and 5 per min. Type III activity appeared as baseline shifts. It was suggested that type I waves represent "mixing"; type II, peristalsis; and type III, a change in muscle tone of the lumen affecting its diameter.

Several questions concerning the EGG must be addressed before it can be used as an indicator of g.i. motility, and, as such, an appropriate feedback variable for the treatment of motion sickness. First, and perhaps foremost, does the surface recording of electrical activity adequately detect and distinguish mechanical events within the stomach and intestine similar to the relationship which exists between mechanical and electrical activity recorded from direct electrode placement on g.i. tissue? Second, is there signifi-

cant interference in the EGG from other bioelectric potentials, such as the EKG and myoelectric activity from respiratory muscles? Third, electrodes are subject to movement artifacts caused by mechanical variations in the skin-electrode interface during movement from activities such as breathing and voluntary changes in position by the subject. How do these movement artifacts affect the EGG? Fourth, can electronic amplifiers be constructed to handle the EGG signals and provide reliable representation of the actual EGG? Fifth, can the electrical signal be readily and reliably associated with a corresponding mechanical event of activity? Sixth, how is the EGG effected by nausea? Seventh, can subjects be taught to modify their electrogastic activity and, if so, will this help them to gain control over their body's response to motion stimuli?

#### IV. Efficacy and Reliability of the EGG.

The relationship between mechanical and electrical events in the gastrointestinal tract has been considered by Stevens and Worrall [1974]. They recorded the EGG from surface electrodes with simultaneous measurements of stomach wall movement via strain gauges in the sedated cat. They reported the results from three of seventeen animals and submitted a three minute record from one animal to

time-series analysis on the computer. Cross-correlation analysis of electrical and mechanical activity indicated a correspondance between these two signals predominately at 4.2 cycles/min. The low frequency range observed for all three animals was between 3.76 and 4.54 cycles/min. In addition, the electrical signal contained activity at higher frequencies concentrated at 9.4, 11.8, 13.9, and 18.6 cycles/min. These higher frequency electrical components did not have a corresponding mechanical equivalent. Although the electrical activity contained frequencies which were not represented in the recording of mechanical activity of the stomach, the higher frequency components corresponded with electrical signals recorded directly from the alimentary canal by others [Stevens, 1974].

Stevens and Worrall concluded that the high degree of correlation, over 80%, between electrical and mechanical activity at the low frequency range was highly significant and indicated the two measurements were correlated. A high correlation between two signals does not necessarily imply a causal relationship, however. These investigators conceded that their results indicated that the mechanical and electrical activities are not exact replicas of each other but are "loose-coupled".

Thus in response to questions 1 and 5 Stevens and Wor-



AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

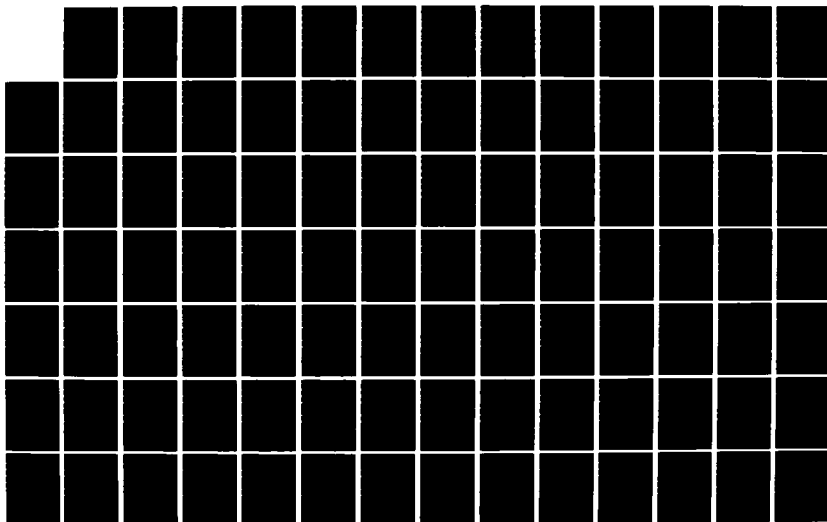
2/13

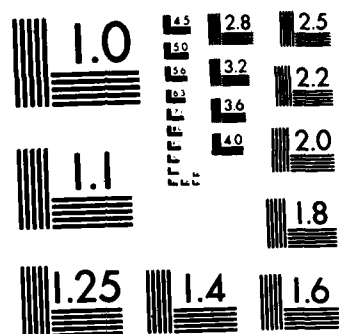
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

rall would reply that surface recorded electrogastric activity contains components correlated with stomach motility but, in addition, reflects the complex nature of electrical activity of the stomach and intestine. Therefore, interpretation of the EGG as stomach motility is not straight forward. These investigators experienced artifacts in their recordings which limited the amount of data which could be analyzed. Since their animals were sedated, it is assumed that voluntary movement did not significantly contribute to these artifacts. Movement from breathing as well as myoelectric interference likely produced these observed artifacts.

The relationship between nausea and EGG activity was investigated by Patterson [1967]. He recorded the EGG and thoracic breathing amplitude from a strain gauge around the chest. He used involuntary head tilts during Coriolis chair rotation to provide motion stimuli to human subjects. His subjects reporting nausea during the head movements had a corresponding increase in the EGG drift and breathing amplitude. (EGG drift was determined from the cumulative sum of the differential deflections recorded every 30 sec.) The group of subjects who were able to tolerate head tilts during rotation with no reports of nausea exhibited little or no increase in either EGG or breathing amplitude. Although Patterson's studies imply that both thoracic breathing am-

plitude and EGG drift correlate with the subjective sensation of nausea, the degree of perceived nausea may not be as well related. The experimental protocol was designed to allow the subject to terminate the experiment when the subject felt that the nausea was no longer tolerable. Patterson compared the EGG and respiratory amplitude signals with the time to termination as designated by the subject. A statistically significant relationship did not exist, which indicated that the subjective awareness of nausea may not be related in a one-to-one manner to thoracic breathing or EGG activity.

Walker and co-workers [1978] looked into the question of voluntary control of the tonic EGG signal. They recorded thoracic respiration, heart rate, abdominal EMG and digital blood flow in addition to EGG. Subjects were asked to voluntarily increase or decrease the tonic EGG activity during a 1.5 min session. The direction of required EGG movement was reversed during the next session with a 25 sec interval between sessions. When analyzed as a group, subjects were able to significantly increase their tonic EGG activity but were unable to reduce it below baseline with statistical significance. Subjects who were successful in controlling their tonic EGG did so with reduced amount of abdominal EMG during both increased and decreased EGG attempts. The unsuccessful subjects exhibited large increases in EMG during

sessions in which an increase in EGG was attempted and decreased EMG during reduction in EGG attempts. The successful subject reported that relaxation and "indifference" were the techniques utilized to control the tonic EGG response during either increase or decrease challenges.

#### V. EGG Recordings.

The electrogastrogram was recorded from two volunteers in order to validate the EGG instrumentation. Electrodes were attached to the abdomen utilizing position 6 from the placement configuration shown in figure III-4 from Earl and Peterson [1983]. The reference electrode was located 1 in. above the navel with the active electrode placed 2 in. above and 2 in. to the subjects' right. The electrode site was prepared by cleaning the skin with an alcohol pad, and standard pregelled silver/silver chloride electrocardiogram electrodes (Instruments for Cardiac Research, Inc.; Catalog #5015) were attached.

The EGG amplifier was provided by the Air Force Institute of Technology [Earl, 1983]. It was designed and constructed as a low-frequency, band-pass, high-gain, differential amplifier. The maximum gain occurred at 0.03 Hz (1.8 cycles/min) with a half-power bandwidth from approximately

0.01 to 0.08 Hz (0.6 to 4.8 cycles/min). The amplifier's gain was adjustable from approximately 98 dB (80,000) to 103 dB (140,000). An offset adjust potentiometer was added to the next to last stage of the amplifier. This offset adjustment was found to be necessary to zero the amplifier's output with shorted inputs. It was not included for the purpose of bucking d.c. potentials present at the input of the amplifier but to compensate for d.c. offsets existing within the operational amplifiers of the final two stages.

Three experiments were performed with subjects sitting or reclined in a comfortable recliner chair. Electrode placement was similar for each experiment as previously described. A record of the EGG from the first experiment is presented in figure 1. This experiment was performed to test the contribution of voluntary abdominal muscle contraction and diaphragmatic breathing on the EGG signal. Abdominal muscle contraction adds a large artifact to the EGG which saturated the amplifier requiring approximately 30 sec to recover. Diaphragmatic breathing produced a smaller artifact which was superimposed on the EGG signal.

Forward body tilts at the waist were performed during experiments 2 and 3. The forward tilt has been used to produce a motion stimulus for a subject spinning in the Coriolis chair. The purpose of these two experiments were to

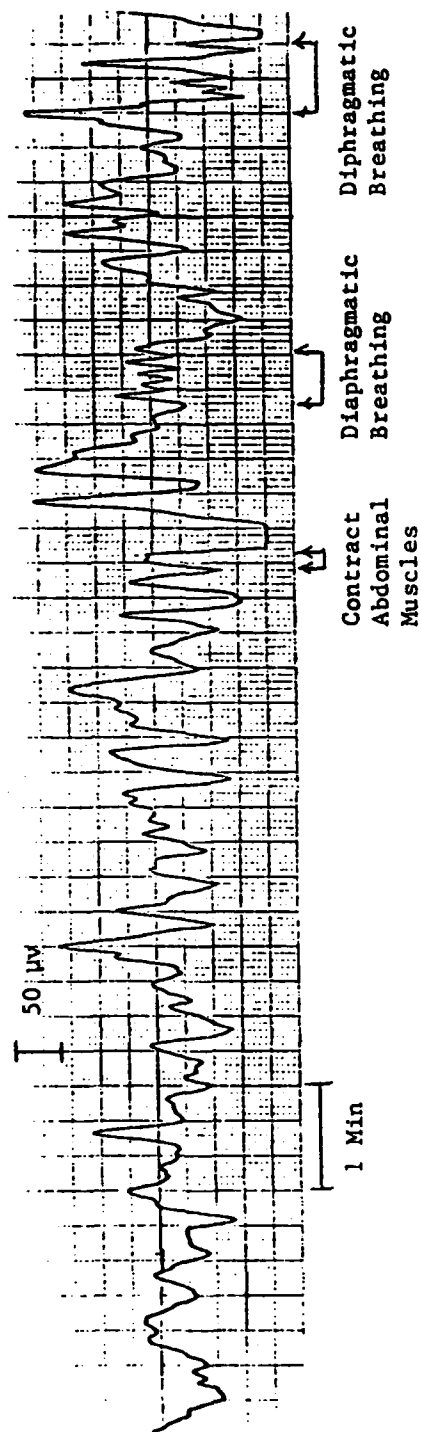


Figure 1: Electrogastrogram EGG record from subject #1 sitting in a chair. During the first 8 minutes the subject was relaxed. He was then instructed to voluntarily contract his abdominal muscles for approximately 10 sec. There was a corresponding large scale deflection in the EGG. Next he breathed using his diaphragm. The resultant EGG was correlated with breathing. A later diaphragmatic breathing epoch appeared to have a breathing artifact superimposed on the EGG signal. Amplifier gain was set at approximately 80,000.

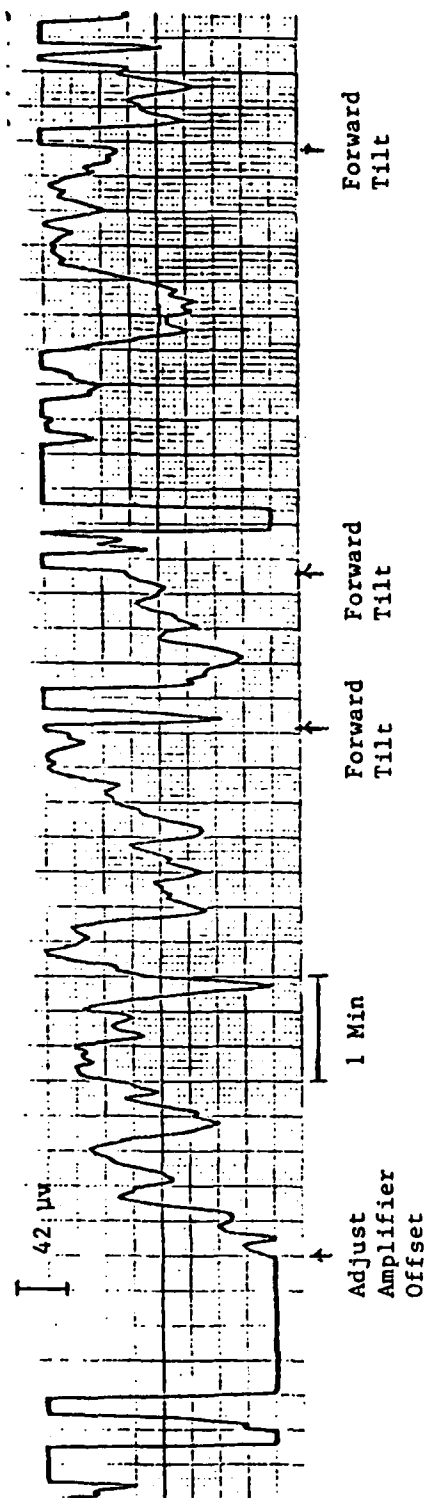


Figure 2: ECG recording from subject #2. This record was taken from the end of an one hour experiment. The subject was sitting upright and relaxed except as noted. During the latter half of this record the subject was asked to bend forward at the waist and return to the upright sitting position. This action required 5 sec or less to complete. The amplifier's output reached its maximum value and recovered within 20 sec during the first procedure. During the second procedure approximately 2 min was required for the amplifier to fully recover. The third forward tilt caused the amplifier to exceed its maximum and recover within 20 sec. At the beginning and again at the end of this record the signal reached the maximum range of the amplifier without a noted movement from the subject. An offset adjustment seemed to compensate for this at the start of the record. Amplifier gain was 95,000.



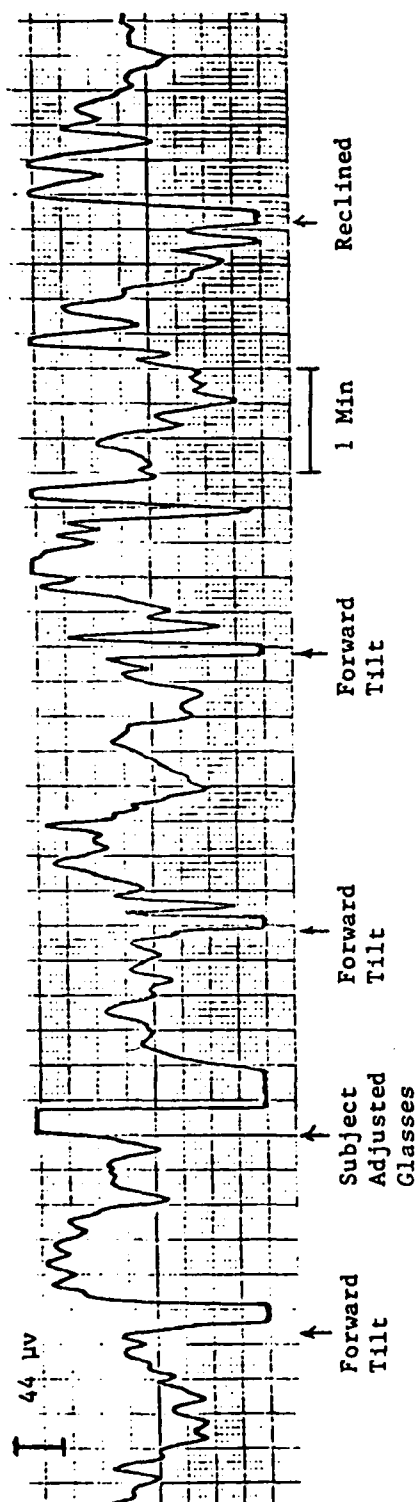


Figure 3: ECG recording from a second experiment for subject #1. The subject was sitting upright in the first part of this record and reclined to the supine position toward the end. Where indicated the subject was instructed to bend forward at the waist and return to the upright sitting position. During the first half of his record the subject voluntarily moved his arm to adjust his glasses. The resultant ECG responded by exceeding the range for the amplifier and required approximately 1 minute to recover. Amplifier gain was 90,000.

determine how the forward tilt effects the EGG in the absence of motion stimulus. Subjects set upright in a stationary chair and tilted the upper body forward and back to the upright sitting position on command. The forward tilt and return to upright took less than 5 sec. In each case the forward tilt produced a large deflection in the EGG recording which exceeded the limits of the amplifier. In many cases the amplifier recovered in 30 sec or less but in a few trials recovery required much longer. As the subject tilts forward the electrode may move in relation to the skin. This movement likely produces a change in the skin-to-electrode potential which would be detected by the EGG amplifier. If the potential is of sufficient magnitude, the front-end amplifier or subsequent stage could be saturated, i.e., driven above its operational limits.

#### VI. Recommendations.

The EGG offers the potential for a direct measurement of gastric motility which could be extremely useful as an index of nausea during motion stimuli. If subjects could be trained to control their electrogastic activity, they should be able to increase their tolerance to motion stimuli. Also, the EGG may offer an independent objective measure of the level of nausea. However, the EGG is not without

its problems. It appears to be highly susceptible to artifacts from movement and interference from other biopotential signals such as abdominal muscle EMG. The EGG contains electrical activity which can not be accounted for by stomach contractions as indicated by the studies of Stevens and Worrall [1974]. The electrical activity uncorrelated with mechanical activity of the stomach may be due to artifacts or it may be from other segments of the alimentary canal. It is important to understand the contributions to the EGG, and it would appear that additional studies are required to better relate g.i. motility with the EGG.

Patterson [1967] found that both thoracic respiratory breathing depth and EGG drift increased with a feeling of nausea. He was however unable to show a correlation between breathing or EGG and the degree of nausea. Although a correlation may be obtained between EGG drift and the degree of nausea with a larger data set, this would not be helpful for evaluation of a individual subject. Therefore other parameters of the EGG should be considered, e.g., a reduction in the 3 cycles/min frequency component.

A question of immediate importance to the application of EGG to biofeedback training for motion stimuli tolerance is the effect of movement artifact on the EGG recording. The test performed in the lab indicated that the forward

body tilt did produce an artifact in the EGG which required from 30 sec or longer to recover. Recovery should be enhanced by redesigning the amplifier by (1) a.c. coupling the front-end stage, and (2) clamping the range of the operational amplifier to prevent saturation.

### References:

- Alvarez, W.C. and L.J. Mahoney. Action currents in stomach and intestine. Am. J. Physiol. 58:476-493, 1922.
- Chaikin, A. Sick in space. Science 84. pp. 51-55, June 1984.
- Cowings, P.S., J. Billingham, and B.W. Toscano. Learned control of multiple autonomic responses to compensate for the debilitating effects of motion sickness. Therapy in Psychosomatic Med. 4: 318-323, 1977.
- Earl, O.A. and C.N. Peterson. Physiological data acquisition system and motion sickness prevention trainer. (Thesis: AFIT/GE/EE/83D-19) Air Force Institute of Technology, Wright-Patterson AFB, Ohio, 1983.
- Goodman, E.N. Improved method of measuring the potential difference across the human gastric membranes and its clinical significance. Surg. Gynecol. Obstet. 75:583-592, 1942.
- Graybiel, A., C.D. Wood, E.F. Miller, and D.B. Cramer. Diagnostic criteria for grading the severity of acute motion sickness. Aerospace Med. 39:453-455, May 1968.
- Graybiel, A. and J.R. Lackner. Motion sickness: acquisition and retention of adaptation effects compared in three motion environments. Aviat. Space Environ. Med. 54: 307-311, 1983.
- Jones, D.R. et al. Control of airsickness through biofeedback-mediated self-regulation of psychophysiological responses. (submitted for publication).
- Joyce, C. Space travel is no joyride. Psychol. Today. pp. 30-37, May 1984.
- Kohl, R.L. and J.L. Homick. Motion sickness: a modulatory role for the central cholinergic nervous system. Neurosci. Biobehav. Rev. 7: 73-85, 1983.
- Levy, R.A., D.R. Jones, and E.H. Carlson. Biofeedback rehabilitation of airsick aircrew. Aviat. Space Environ. Med. 52(2): 118-121, 1981.
- Oman, C.M. Space motion sickness and vestibular experiments in spacelab. SAE Technical Paper Series (820833), Twelfth Intersociety Conference on Environmental Systems, San Diego, CA, July 19-21, 1982.
- Patterson, M.M. Changes in the electrogastrogram and in breathing amplitude as indicators of nausea. Tech. Rep. 18, Contract Nonr 908-15, Office of Naval Research, June 1967.

Russell, R.W. and R.M. Stern. Gastric motility: the electrogastrogram. In: A manual of psychophysiological methods. P.H. Venables and I. Martin (Eds.). Amsterdam: North-Holland Pub. Co., 1967.

Stevens, J.K. and N. Worrall. External recording of gastric activity: the electrogastrogram. *Physiol. Psych.* 2:175-180, 1974.

Walker, B.B., C.A. Lawton, and C.A. Sandman. Voluntary control of electrogastric activity. *Psychosomatic Med.* 40:610-619, 1978.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

FAR-INFRARED ABSORPTION PROFILES FOR SHALLOW DONORS

IN GaAs-GaAlAs QUANTUM WELL STRUCTURES

Prepared by:	Dr. Ronald L. Greene
Academic Rank:	Associate Professor
Department and	Department of Physics
University:	University of New Orleans
Research Location:	Air Force Avionics Laboratory Electronic Research Branch
USAF Research	Dr. Krishan K. Bajaj
Date:	July 25, 1984
Contract No:	F49620-82-C-0035

FAR-INFRARED ABSORPTION PROFILES FOR SHALLOW DONORS

IN GaAs-GaAlAs QUANTUM WELL STRUCTURES

by

Ronald L. Greene

ABSTRACT

Previous variational calculations of the ground and 2p-like excited states of a shallow donor in a quantum well are extended to include the effects of an applied magnetic field and arbitrary donor position. The extended wave functions are then used in a theory for the absorption profile of shallow donor transitions in quantum wells. Absorption profiles for several donor impurity distributions are calculated and compared to recent far-infrared absorption experiments. The comparison with experiment is ambiguous, but the theory does suggest that the experimental samples may possess thin donor impurity layers at the interfaces between semiconductors.



#### ACKNOWLEDGMENTS

I would like to thank Dr. Krishan Bajaj of the Electronic Research Branch for stimulating and helpful discussions throughout the ten weeks of research summarized in this report. I would also like to acknowledge the interesting lunches and conversation supplied by Dr. George Norris, Dr. Bill Iheis, and Lt. Larry Kapitan that kept me going through the summer.

Finally, I thank the Air Force Systems Command, the Air Force Office of Scientific Research, the Air Force Avionics Laboratory (particularly the Electronic Research Branch), and the Southeastern Center for Electrical Engineering Education for providing me with the opportunity to conduct this research.

## I. INTRODUCTION

Current epitaxial growth techniques such as molecular beam epitaxy and metal-organic chemical-vapor deposition make it possible to grow systems of alternate layers of different semiconductors (usually referred to as heterostructures) or layers of a single semiconductor with different doping properties. Such systems with large-scale periodicity (typically 50-500 Å) along the growth direction are known as superlattices. Perhaps the most commonly studied heterostructure consists of alternating layers of GaAs and  $\text{Ga}_{1-x}\text{Al}_x\text{As}$ , where the Al fraction  $x \leq 0.4$ . In this paper we shall confine our attention to such systems.

The band gap of  $\text{Ga}_{1-x}\text{Al}_x\text{As}$  varies with  $x$ , and is larger than that of GaAs. When it is grown in alternating layers with GaAs, discontinuities in the conduction and valence band edges occur at the interfaces. For sharp interfaces the potential felt by an electron moving in the conduction band is that of a one-dimensional array of square wells and barriers, with the wells formed in the GaAs and the barriers in the  $\text{Ga}_{1-x}\text{Al}_x\text{As}$ . The presence of these quantum wells causes the normal three-dimensional band structure to be split into subbands along the  $k_z$  direction. In the effective mass approximation the subband edges correspond to the energy levels of an effective mass electron moving in the one-dimensional potential.

Studies of shallow donor states in bulk semiconductors have yielded considerable information about the host semiconductor, so that it is not surprising that similar studies have been made of shallow donors in heterostructures. The energy levels associated with such shallow states in a GaAs- $\text{Ga}_{1-x}\text{Al}_x\text{As}$  quantum well are found near the conduction subband

edges. The Coulomb binding energy has been the focus of most of the theoretical work to date.<sup>1-2</sup>

Recently far-infrared absorption experiments have begun to measure the shallow donor 1s-2p transitions in GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As quantum wells. At the present time the absorption profiles are rather poor, but several peaks have been seen. There is now a need for accurate theoretical absorption profiles for comparison. In principle, such profiles can be combined with experimental profiles to extract the distribution of impurities in a given sample. Prior to this work there has been one calculation of the absorption line shape for the 1s to 2p shallow donor transition in these heterostructures.<sup>3</sup> There are several problems with the calculation, however, in terms of comparison to experiment. First, the model chosen was that of a quantum well with infinitely high barriers. This has been shown to be not a very good approximation for narrow wells ( $L \lesssim 50\text{\AA}$ ). A second problem is that the line shape was calculated for zero external magnetic field. The experimental data, on the other hand, is much better for moderate size magnetic fields. However, the biggest problem with the previous calculation is that the authors arbitrarily considered only donors within the GaAs well in their calculation of the line shape. Tanaka, *et al.*<sup>4</sup> have shown that the binding of electrons within the well to donors outside it can be appreciable. This, plus the fact that the density of states continues to increase for donors outside the well means that they should not be neglected. In fact, as will be pointed out later, neglecting such impurities leads to the prediction of a spurious second peak in the absorption profile.

This problem of the calculation of absorption profiles of shallow

donor transitions in quantum well systems is an obvious one for me to work on. Not only is the work of importance in the understanding of the physics of semiconductor heterostructures, but my experience with several aspects of the problem make me especially well qualified to tackle it. Several collaborators and I have previously published the results of work on far-infrared shallow donor spectra in GaAs, on the effects of magnetic fields on shallow donors, and on the effects of quantum wells upon the energy levels of shallow donors in GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As materials. I also have extensive experience in the calculation of line shapes for hydrogenic atoms in plasmas.

## II. OBJECTIVES OF THE RESEARCH EFFORT

The specific objectives of my research were all aimed at producing theoretical absorption profiles of 1s-2p shallow donor transitions in GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As quantum well systems. They are listed below:

- 1) Apply the theory for the absorption profiles to shallow donors in quantum wells with an external magnetic field.
- 2) Obtain accurate ground and excited state variational wave functions as a function of donor position for several values of magnetic field and well width.
- 3) Evaluate the transition line strengths using these wave functions.
- 4) Calculate the line shape for a variety of impurity distributions.
- 5) Compare the results with experiment.

### III. THEORY OF THE ABSORPTION PROFILE

The theory of absorption of light by atomic systems is well known, and can be found in standard texts on quantum mechanics (See, for example, Reference 5.). We will only sketch it here, and indicate specifically how it can be applied to the problem of shallow donor absorption. We first assume that our donors are sufficiently well separated that they do not interact with each other. Since the effective Bohr radius of a shallow donor in GaAs is about 100 Å, and our donors will be constricted still further by the applied magnetic field and the quantum well, this is not a very restrictive approximation. We also assume that the experiments will be performed at very low temperature so that all the donors are initially in their ground states. We choose our coordinate system such that the z-axis is directed along the growth direction for the heterostructure, and consider the incident light polarized in the x-direction with intensity  $I_0$ .

The probability per unit time for an electric dipole transition from an initial state  $i$  to a final state  $f$  is given by

$$P_{i \rightarrow f} = \frac{4\pi^2 e^2 \hbar^2}{m^* E_{fi}^2 c} I_0 |\langle f | \nabla_x | i \rangle|^2 \delta(\hbar\omega - E_{fi}), \quad (1)$$

where  $\omega$  is the angular frequency of the incident light and  $E_{fi}$  is the energy difference between the final state and the initial state ---  $E_{fi} = E_f - E_i$ . In the effective mass approximation the mass in Equation (1) is the effective mass of the donor electron.<sup>6</sup> If we multiply this expression by  $E_{fi}$  and sum over final states we obtain the rate of energy loss from the field caused by absorption as a donor electron is excited from an initial state  $i$  to any other state. Then if we multiply the resulting expression by  $P_i$ , the probability that the donor electron is

initially in the state  $i$ , and sum over all initial states, we get the following expression for the power absorbed from the incident radiation:

$$P(\omega) = \frac{4\pi^2 e^2 I_0}{m^2 \omega c} \sum_{i,f} |\langle f | \nabla_x | i \rangle|^2 \rho_i \delta(\omega - E_{fi}/\hbar). \quad (2)$$

Kohn<sup>6</sup> has shown that within the effective mass approximation

$$|\langle f | \nabla_x | i \rangle|^2 = \frac{m^2 E_{fi}^2}{\hbar^2} |\langle f | x | i \rangle|^2. \quad (3)$$

Substituting this into Equation (2) yields the following expression for the absorbed power:

$$P(\omega) = \frac{4\pi^2 e^2 \omega I_0}{\hbar^2 c} \sum_{i,f} |\langle f | x | i \rangle|^2 \rho_i \delta(\omega - E_{fi}/\hbar). \quad (4)$$

In the presence of an applied magnetic field,  $-\hbar \nabla^2$  should be replaced by  $-\hbar \nabla^2 - (e/c) \vec{A}^2$ , where  $\vec{A}$  is the magnetic vector potential. However, the relationship in Equation (3) is changed with the result that Equation (4) is still valid.<sup>7</sup>

For a bulk semiconductor the sum over  $i$  reduces to a single term since the ground state energies of all isolated donors are the same. However, for a quantum well system the ground state energy of a given donor depends on its location relative to the quantum well. The sum over  $i$ , then, is actually a sum over all the possible positions of the donor impurity.

In this work we are interested in the shape of the absorption profile, rather than its absolute expression. For this reason we define a function  $I(\omega)$  as follows:

$$I(\omega) \equiv \sum_{i,f} |\langle f | x | i \rangle|^2 \rho_i \delta(\omega - E_{fi}/\hbar). \quad (5)$$

$I(\omega)$  is usually called the absorption line shape because it contains most of the frequency dependence of the power spectrum of Equation (4). It is closely related to the imaginary part of the complex dielectric function.

The sum over donor positions in Equation (5) is a sum over the Ga and Al sites in the GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As heterostructure. However, since the size of the donor is many times larger than the lattice spacing, the sum may be approximated by an integral with little error. The energy of a donor is dependent only upon its z-coordinate ( $z_i$ ), so we can replace the sum over  $i$  by an integral over  $z_i$ . The probability  $p_i$  in this case becomes the linear density distribution of donors,  $p(z_i)$ .

As we have noted, the transition energy  $E_{fi}$  is a function of the impurity position  $z_i$ . We may formally invert the relationship to obtain  $z_i$  as a function of angular frequency  $\omega_{fi} = E_{fi}/\hbar$ . This allows us to perform the integration to obtain

$$\begin{aligned} I(\omega) &= \sum_f \int_{-\infty}^{\infty} dz_i p(z_i) |\langle f|x|i \rangle|^2 \delta(\omega - \omega_{fi}) \\ &= 2 \sum_f \left| \frac{d\omega_{fi}}{dz_i} \right|^{-1} |\langle f|x|i \rangle|^2 p(z_i) . \end{aligned} \quad (6)$$

Because of the cylindrical symmetry of the problem, the initial (ground) states are states with zero angular momentum projection along the z-axis (i.e.,  $m=0$  states as discussed below). The matrix elements of the  $x$  operator will then vanish for all states except those for which  $m=\pm 1$ . At this point we assume that the frequency of the incident radiation is such that only the 2p-like states contribute significantly to the line shape in the region of interest. This point will be discussed later. This approximation reduces the sum over final states to just two terms, the

$2p_{\pm}$ -like states, which from now on we will refer to as  $2p_{\pm}$ . We defer a description of these states until the next section. Furthermore, the application of a magnetic field will cause the  $2p_{+}$  and  $2p_{-}$  states to split, so that possibly only one term will be important for the line shape over a given range of frequency.

#### IV. SHALLOW DONOR WAVE FUNCTIONS

In order for us to obtain a numerical absorption profile it is necessary that we have available accurate representations of the ground and  $2p_{\pm}$  states for the shallow donor in a quantum well. This problem, with an applied magnetic field along the growth direction, has been examined by Greene and Bajaj.<sup>8</sup> We will sketch the analysis here; see Reference 8 for details.

In the effective mass approximation the Hamiltonian for the shallow donor electron may be written in the following form:

$$H = -\frac{1}{m^*} \nabla^2 - \frac{2}{r} + \gamma L_z + \frac{1}{4} \gamma^2 \rho^2 + V_w(z). \quad (7)$$

This equation has been written in dimensionless form. The unit of energy is the effective Rydberg and the unit of length is the effective Bohr radius --- both determined using the parameters of bulk GaAs. (These quantities are 5.83 meV and 98.7 Å, respectively.) The quantity  $m^*$  in the above expression is the effective mass of the electron in units of the bulk GaAs effective mass. Inside the well  $m^*=1$ ; in the  $\text{Ga}_{1-x}\text{Al}_x\text{As}$  barrier material it is assumed to vary with  $x$  as<sup>2</sup>

$$m^* = 1 + 1.24x. \quad (8)$$

In Equation (7)  $L_z$  is the z-component of angular momentum (in units of  $\hbar$ ),



$\rho = [x^2 + y^2]^{1/2}$ , and  $V$  is a dimensionless measure of the magnetic field, the ratio of the lowest Landau energy of a "free" effective mass electron to the effective Rydberg. The quantum well potential  $V_w(z)$  is assumed to be a simple square well. Neighboring wells are considered to be sufficiently far away that there is virtually no penetration of the donor electron wave function into another well. This requires the  $\text{Ga}_{1-x}\text{Al}_x\text{As}$  barrier to be wide ( $\geq 300 \text{ \AA}$ ).

The variational wave functions that we use to evaluate the line shape function are of the same form as we used to find the binding energies of a shallow donor in a well with an applied magnetic field. Since the Hamiltonian in Equation (7) is cylindrically symmetric, the component of angular momentum along the  $z$  axis is conserved, so that the  $\phi$  dependence of the wave function has the form  $\exp(i m \phi)$ . Furthermore, since for most well widths the Coulomb contribution to the energy is considerably smaller than the contribution from the square well potential, it is helpful to explicitly factor the solution for the ground state of the one-dimensional square well problem out of our variational wave function. We label this square well function  $f(z)$ . Its explicit form can be found in Reference 8 or any general quantum mechanics text.<sup>5</sup> The form for the effective mass envelope function that we have used is then given by

$$\psi(\vec{r}) = \rho^{|m|} e^{i m \phi} f(z) \sum_{j,k} A_{jk} G_{jk}(\rho, z), \quad (9)$$

where the basis functions  $G_{jk}(\rho, z - z_j)$  are taken to be the product of Gaussian functions in  $\rho$  and  $z - z_j$ .

$$G_{jk}(\rho, z - z_j) = e^{-(\alpha_j + \beta) \rho^2} e^{-\alpha_n (z - z_j)^2} \quad (10)$$

The parameters  $A_{jk}$ ,  $\alpha_j$ , and  $\beta$  were determined variationally. Thirteen terms were used in the expansion of Equation (9) for both the ground and excited states.

## V. RESULTS AND DISCUSSION

As we indicated earlier, the major broadening effect for shallow donor transitions in these heterostructures is the position of the donor with respect to the well. This is clearly indicated in Figure 1, which shows the dependence of the  $1s-2p_-$  transition energy as a function of the position of the donor. The magnetic field chosen for Figure 1 corresponds to  $\gamma = 1$ , and the energy units are effective Rydbergs (Ryd). (The  $1s-2p_+$  energy can be obtained by adding  $2\gamma$  to the transition energies in this, and the succeeding, figures.) The donor position,  $z_i$ , is given as a fraction of the well width  $L$ , with  $z_i=0$  being the center of the well. As can be seen in the figure, the  $1s-2p_-$  energy decreases monotonically as the donor is moved from the center of the well to the edge ( $z_i=L/2$ ) and on into the barrier material. Note the wide range ( $\sim 1$  Ryd) of the transition energy for donors at different well positions.

As seen in Equation (6), the derivative of the transition energy with respect to the donor position is an important factor in determining the absorption profile. This derivative vanishes at  $z_i=0$ , causing the ideal profile to have an infinite spike at the energy corresponding to  $z_i=0$ . Other broadening mechanisms neglected in Equation (6) will remove the singularity, but there will be a peak characteristic of donors located at the center of the well.

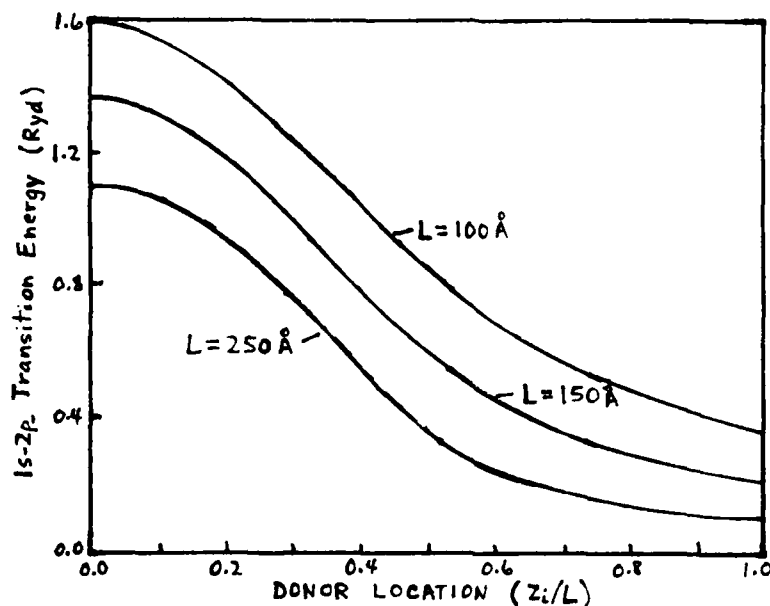


Figure 1. Variation of the 1s-2p<sub>z</sub> transition energy with donor position.

the magnetic field parameter  $\gamma = 1$ .

The solid line of Figure 2 shows the absorption profile for a uniform impurity distribution versus transition energy over a range of about 1 effective Rydberg. As in Figure 1, the magnetic field value is  $\gamma = 1$ , while the well width is 150 Å. The points shown on the solid curve correspond to equally-spaced  $z_i$  values, each separated from the two neighboring points by  $0.05L$ , or 7.5 Å. The edge position,  $z_i = L/2$  is indicated on the figure, while the center position is at the right side, at the location of the absorption spike predicted in the preceding paragraph. Note that there is only the single spike in the uniform donor distribution curve. The absorption continues to increase for still smaller energy values than shown in the figure. This is due to the fact

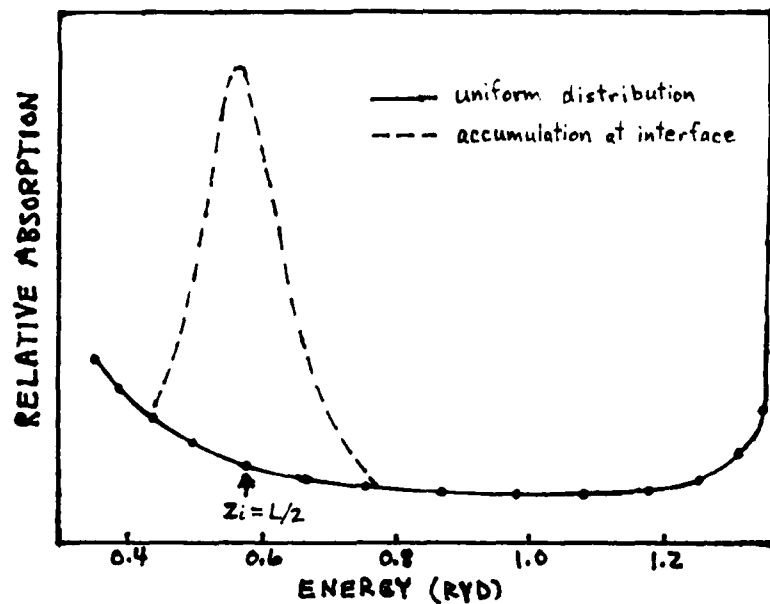


Figure 2. Absorption profile for a uniform distribution of donor impurities and for an accumulation of donors at the interface.  $L=150 \text{ \AA}$  and  $\gamma=1$ .

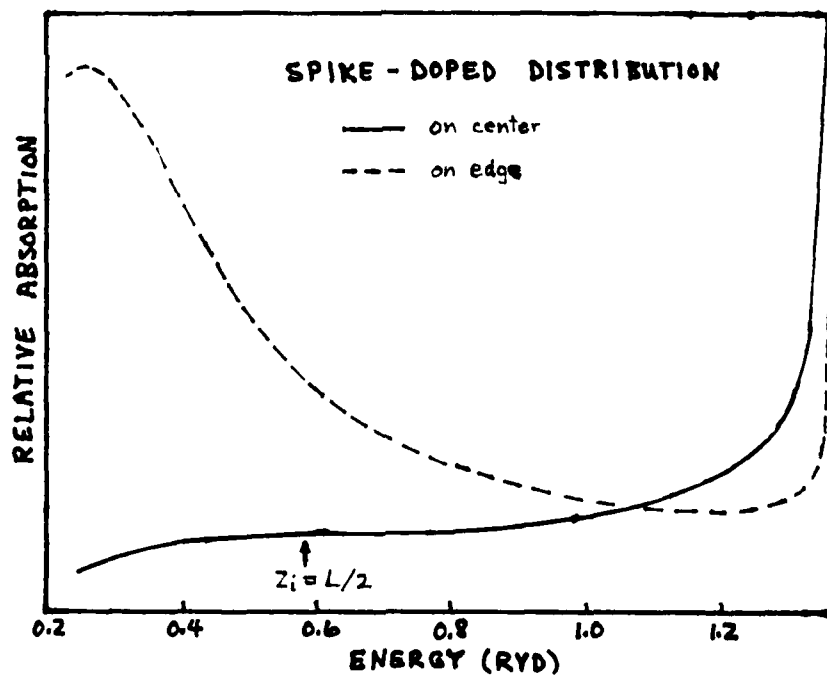


Figure 3. Absorption profile for spike doping at the center and at the edge of a GaAs quantum well.  $L=150 \text{ \AA}$  and  $\gamma=1$ .

that the derivative of the transition energy with respect to  $z_i$  becomes smaller with increasing  $z_i$ . (See Figure 1.) This increase will not continue indefinitely in a real heterostructure with finite-width barriers because as a donor moves further into a barrier away from one quantum well, it gets closer to a neighboring well. Thus a finite-width barrier results in some maximum value for  $z_i$ , and for uniform doping should exhibit a peak characteristic of donors at the center each barrier.

The dashed curve in Figure 2 illustrates the absorption profile with a 10 Å wide Gaussian donor impurity distribution superimposed on a uniform background distribution. The Gaussian distribution is centered at the interface between the GaAs and  $\text{Ga}_{1-x}\text{Al}_x\text{As}$ , and has a peak value of five times that of the background. The presence of a local impurity concentration at the interfaces between semiconductors is suggested by the photoluminescence spectra of Shanabrook and Comas,<sup>9</sup> who observed two donor-related peaks at energies in approximate agreement with the on-center and on-edge shallow donor calculations of Mailhot, et al.<sup>2</sup> (The theoretical prediction of Bastard<sup>1</sup> that there should be a second peak in the spectrum of a uniform donor distribution is in error. His prediction applies to heterostructures whose donor impurities are found only in the well material.) Figure 2 thus shows that if the donor impurity distribution is uniform, there should be only one peak associated with the  $1s-2p_z$  transition. A second peak can be caused by a local concentration of impurities as might be found at the interfaces between the two semiconductors.

Figure 3 shows the effect of a "spike-doped" donor distribution upon

the absorption profile. Spike doping has been employed by Shanabrook and collaborators in several experimental studies of shallow donors in GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As heterostructures. The profile in Figure 3 is for a 150 Å wide quantum well, with 50 Å wide Gaussian doping spikes. Any background impurity distribution is assumed negligible. The applied magnetic field again corresponds to  $\gamma = 1$ . Two cases are presented in the figure. The solid line represents the case in which the peak of the spike occurs at the center of the well, while the dashed curve represents the results for doping spikes whose peaks are at the interfaces between the two semiconductors. Not surprisingly, the on-center case yields an absorption profile with a single peak whose maximum transition energy is that of a donor located at the center of the well.

The on-edge case is more complicated. There is still a spike associated with donors at the center of the well. However, its presence may be misleading. The amount of absorption seen in a real profile is proportional to the area under the curve in any given energy range. The area under the central spike is relatively much smaller in this case than it is for either the uniform distribution or the on-center spike doping. Another interesting point about the dashed curve of Figure 3 is that the low energy peak of the absorption does not occur at the same place as the peak of the doping spike. This is because the doping spike is very broad and is multiplied by a factor that increases with smaller energies --- the inverse of the derivative of the transition energy with respect to donor position  $z_1$ . The combination causes the peak to be shifted to lower energies than would be expected for a very narrow distribution centered at the interface (as in the dashed curve of Figure 2).

Recently Jarosik, et al.<sup>10</sup> have obtained data from far-infrared absorption experiments on shallow donors in spike-doped GaAs-Ga<sub>1-x</sub>Al<sub>x</sub>As quantum well structures. In one of the samples, for example, with magnetic field strengths in the range  $0.5 \leq V \leq 1.5$  they observe three peaks in a broad absorption feature approximately centered around the bulk GaAs 1s-2p<sub>+</sub> transition. This particular sample is doped with a  $\sim 50$  Å wide spike at the center of the well. Let us assume that, in addition to the doped spike, their heterostructures contain a narrow concentration of impurities at the interface. As mentioned earlier, this assumption is consistent with their photoluminescence data on some of the same samples. The two theoretical peaks corresponding to donors at the center and at the edges of the 150 Å wide GaAs well fall rather close to the two higher energy peaks seen by Jarosik, et al., so it is tempting to identify these experimental peaks with our theoretical ones. Although the third (low energy) peak is at about the right energy for the peak corresponding to donors at the center of the barrier layer (which was not shown in Figures 1-3), the much lower donor concentration in the barrier makes it unlikely that such a peak would appear in the experimental profiles.

Moreover, photoluminescence<sup>9</sup> and Raman scattering<sup>11</sup> experiments suggest that the theoretical values for the binding energy of the ground state of the shallow donor at the center of a quantum well are too large by a few meV. This also appears to be the case for the zero magnetic field far-infrared absorption measurements. If there is a systematic overestimation of the binding energy on the part of the theoretical calculations, then the two lower experimental peaks may correspond to our on-center and on-edge 1s-2p<sub>+</sub> transition energies. The third (high

energy) experimental peak could then be due to higher shallow donor transitions, such as a  $1s-3p_+$  transition.

## VI. RECOMMENDATIONS

As seen above, the comparison between theory and experiment is ambiguous. Far-infrared data has only recently been obtained, so there is need for improvement in the data as well as for more data. For example, it would be helpful if the  $1s-2p_-$  transition could be measured as consistently as the  $1s-2p_+$ . Also, additional studies with varied doping profiles and particularly wider  $Ga_{1-x}Al_xAs$  barriers would be helpful toward identifying the three peaks in the absorption profile.

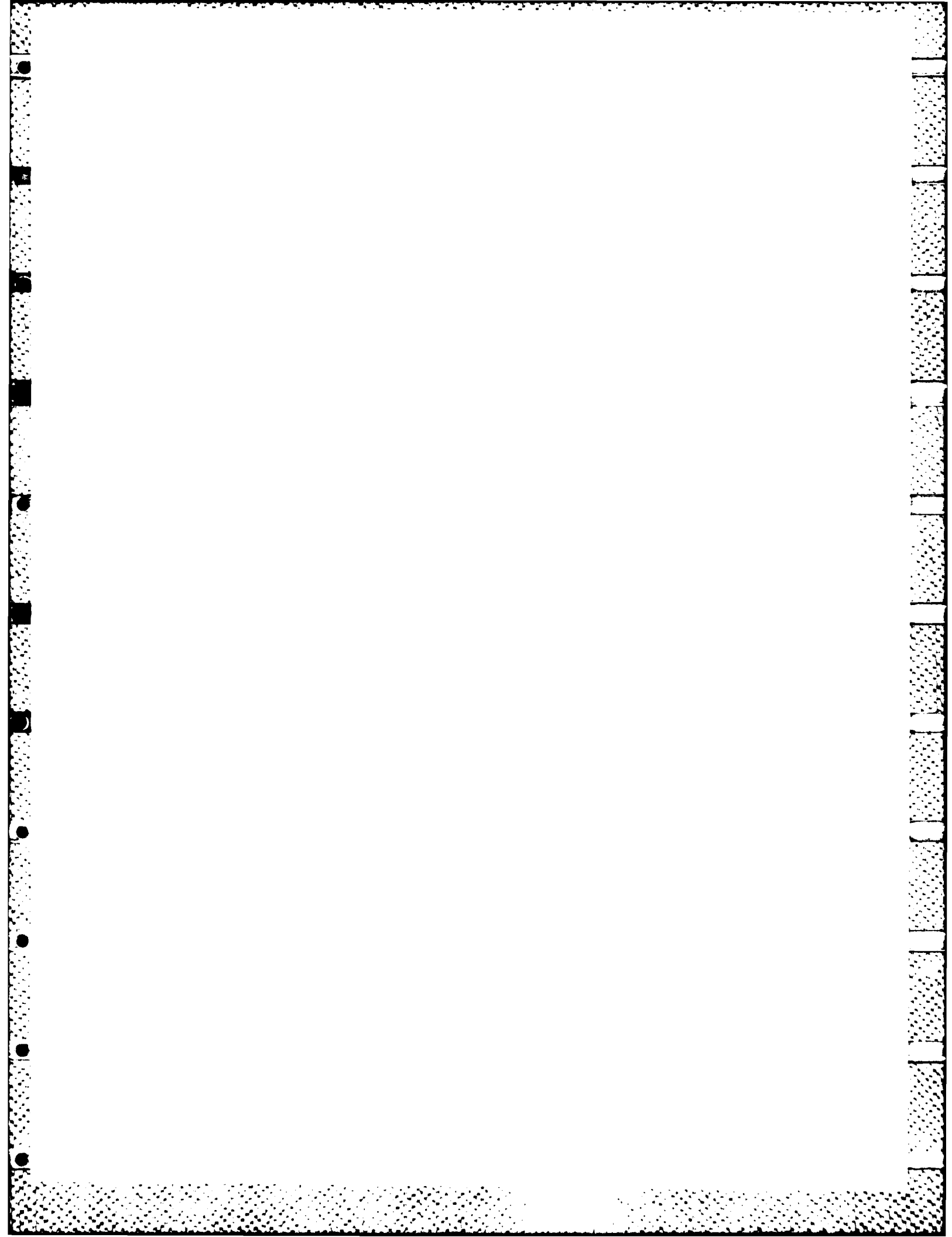
On the theoretical side, there are two primary improvements that can and should be made to the calculated profiles. First, better account should be made of the finite width of the barrier material. Not only would doing so reduce the transition energies somewhat, but more importantly, it would determine the location and the strength of the peak due to donors near the center of the barrier material which we neglected in this work. It would be worthwhile to know whether this peak is one of the ones seen in the far-infrared absorption experiments.

The other improvement that should be made is the inclusion of  $1s-3p$  transitions in the line shape calculation. Determining the location of the  $1s-3p$  peak due to centrally located donors is important since it is a candidate for one of the experimentally observed peaks. Furthermore, it is likely that such transitions from off-center donors will overlap the  $1s-2p$  profile, possibly altering its shape.



#### REFERENCES

1. G. Bastard, "Hydrogenic Impurity States in a Quantum Well," Phys. Rev. B, Vol. 24, pp. 4714-4722, 1981.
2. C. Mailhot, Yia-Chung Chang, and T. C. McGill, "Energy Spectra of Donors in  $\text{GaAs-Ga}_{1-x}\text{Al}_x\text{As}$  Quantum Well Structures in the Effective Mass Approximation," Phys. Rev. B, Vol. 26, pp. 4449-4457, 1982.
3. G. Bastard, E. E. Mendez, L. L. Chang, and L. Esaki, "Far Infrared Impurity Absorption in a Quantum Well," Solid State Commun., Vol. 45, pp. 367-369, 1983.
4. Kazuyoshi Tanaka, Masataka-Nagaoka, and Tokio Yamabe, "Binding Energy of the Impurity Level in the  $\text{Ga}_{1-x}\text{Al}_x\text{As-GaAs-Ga}_{1-y}\text{Al}_y\text{As}$  Superlattice," Phys. Rev. B, Vol. 28, pp. 7068-7074, 1983.
5. Leonard I. Schiff, Quantum Mechanics, 3rd Ed. (McGraw-Hill Book Company, New York, 1968), pp. 402-405.
6. W. Kohn, "Shallow Impurity States in Si and Ge," in Solid State Physics, Vol. 5, ed. by F. Seitz and D. Turnbull (Academic Press, New York, 1957), pp. 283-285.
7. C. Aldrich and R. L. Greene, "Hydrogen-Like Systems in Arbitrary Magnetic Fields," phys. stat. sol. (b), Vol. 93, pp. 343-350, 1979.
8. Ronald L. Greene and K. K. Bajaj, "Effect of Magnetic Field on the Energy Levels of a Hydrogenic Impurity Center in  $\text{GaAs-GaAlAs}$  Quantum Well Structures," submitted for publication.
9. B. V. Shanabrook and J. Comas, "Photoluminescence from 'Spike Doped' Hydrogenic Donors in  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As-GaAs}$  Quantum Wells," in Proceedings of the 5th International Conference on Electronic Properties of Two-Dimensional Systems, to be published in Surface Science.
10. N. C. Jarosik, B. D. McCombe, B. V. Shanabrook, J. Comas, and G. Wicks, to be published.
11. B. V. Shanabrook, J. Comas, T. A. Ferry, and R. Merlin, "Raman Scattering from Electrons Bound to Shallow Donors in  $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$  Quantum-Well Structures," Phys. Rev. B, 7096-7098, 1984.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DIGITAL SIGNAL PROCESSING APPROACHES FOR

ANALYSIS AND EVALUATION OF COMMUNICATION SYSTEMS

Prepared by:	Paul B. Griesacker, Ph.D.
Academic Rank:	Professor
Department and University:	Department of Physics Gannon University
Research Location:	Air Force Wright Aeronautical Laboratory, Systems Avionics Division, Information Transmission Branch, AFJAL/AAAI
USAF Research:	Mrs. Diane E. Summers
Date:	August 8, 1984
Contract No:	F49620-82-C-0035

DIGITAL SIGNAL PROCESSING APPROACHES FOR  
ANALYSIS AND EVALUATION OF COMMUNICATION SYSTEMS

by

Paul B. Griesacker, Ph.D.

ABSTRACT

Current approaches to analysis and evaluation of Communication, Navigation, and Identification systems are discussed with recommendations for AAAI-4 in preparation for the delivery of the Integrated Electromagnetic System Simulator (IESS). Digital signal processing approaches are discussed including benefits of the dual channel FFT analyzer in the presence of noise. Digital system simulations are examined including the need for introducing the RF path by analog methods. Jamming capabilities for IESS are recommended making maximum utilization of hardware already available to AAAI-4.

## I. INTRODUCTION

The Air Force has recognized the need to consolidate the many varied communication, navigation and identification (CNI) systems carried on modern aircraft. A solution to this problem of redundancy of function and incompatibility of hardware is the Integrated Communication, Navigation and Identification Avionics (ICNIA) system<sup>1</sup>. With this innovation, all aircraft will carry a single CNI system. The ICNIA system will consist of a minimum amount of modular hardware and a maximum of software. There will be some deliberate redundancy in the hardware so that a failure in one of the subsystems can be channeled through compatible hardware of another subsystem. This software driven CNI system will also have self diagnostic capabilities. It is apparent that ICNIA will be a complicated system and the hardware will include new technologies. The evaluation of such a complicated and interrelated set of subsystems will be a monumental task especially for testing performance in hostile atmospheres containing noise, fade, jam, multipath, etc., all in real time.

An answer to the ICNIA test and evaluation question is the Integrated Electromagnetic System Simulator (IESS)<sup>2</sup>. IESS is itself somewhat of a software driven communication system. It contains no receiver capability excepting that of the imbedded ICNIA under test. IESS has some, but not all the transmitter capabilities of the CNI services in ICNIA; nor does it have the capability to generate jam. IESS is designed to run software driven real time CNI scenarios with computer stored and analyzed results. IESS itself must be evaluated as an adequate test facility for ICNIA. It will be necessary to begin the

evaluation process early on in the schedule, since the new technology hardware should be evaluated as it evolves so that positive feedback can be given to the contractors during the development stages.

A systematic approach to the evaluation process is necessary to keep the effort manageable and consistent. The linear systems approach is such a method. A linear time invariant system can be characterized<sup>3,4,5</sup> by studying the output of the system relative to specified inputs, and great convenience is introduced if this is done in the frequency domain. Even if a system is not completely linear much information can be gained by finding the gain and phase shift at frequencies of interest. This also, is most conveniently done in the frequency domain. This input/output comparison can be applied to the complete system, i.e. compare the input of the transmitter to the output of the receiver; or it can be applied to one of the subsystems, i.e. one of the amplifiers. Even individual components can be examined in this manner.

This approach to system evaluation and analysis will be discussed in greater detail later in this report. The author has had previous experience in this field<sup>6,7</sup>, particularly in the frequency or Fourier domain.

## II. OBJECTIVES OF THE RESEARCH EFFORT

It was agreed during the pre-summer visit that the effort for this summer research period would be the evaluation of approaches for accomplishing signal processing functions in the area of communications relative to the mission of AAAI. These evaluations should lead to recommendations concerning: mission objectives, hardware specifications

and software approaches. This effort would be accomplished by making a survey of existing industrial and governmental approaches, including an attempt to make optimal usage of existing hardware and software presently available in AAAI.

To a large extent all the above objectives have been accomplished, at least as far as the recommendations are concerned. Although carrying out these recommendations has begun under this program, in several instances a continued effort will be required to accomplish the desired results. Some of these ongoing efforts are being included in a proposal to be submitted under the Research Initiation Program sponsored by AFOSR.

### III. SYSTEMS ANALYSIS

Herein are presented some ideas to be considered in preparation for the installation of the IESS system at WPAFB in 1988. First a general and comprehensive overview of some CNI system evaluation approaches will be reviewed; and finally some specific approaches that have the potential for immediate implementation in AAAI-4 will be presented.

A generalized CNI system can be visualized as in Figure 1. Here a modular block diagram shows the transmitter and receiver as complimentary sets of subsystems (modules) each of which perform some specific operation or process on the signal it receives at its input port. The result of this specific operation will be carried on the signal at the output port of the module. A transmitter accepts an information signal (voice or data) at its input transducer (microphone or data bus) and operates on this signal with several subsystem functions as shown. The antenna itself can be represented by one or

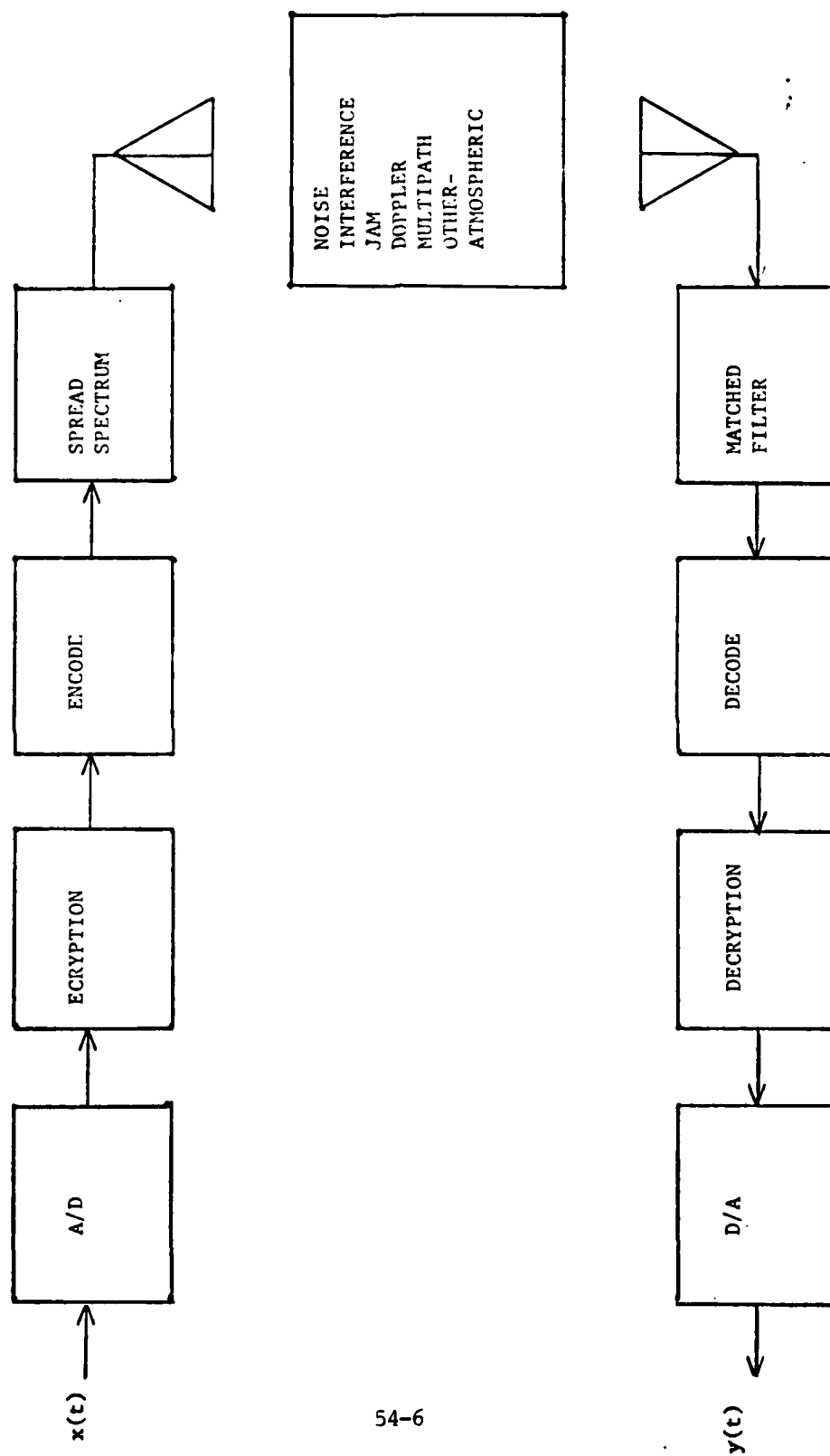


FIGURE 1. GENERALIZED CNI SYSTEM



more modules depending on its sophistication, for instance a simple dipole as opposed to a phased array. Interfacing problems have been ignored for simplicity, but they could be represented as one or more modules in the final analysis. The radio frequency (RF) path can also be represented by several separate modules that operate on the radiated waveform. Here is where jam and atmospheric effects can be taken into account.

The receiver also has an antenna which would have a modular representation complimentary to the transmitting antenna system. The demodulating subsystems of the receiver, most of which perform an inverse operation of the corresponding modulation block in the transmitter system, also can be modularized and again interfacing effects can be represented by one or more system blocks.

This generalized overview of CNI systems, though simplistic in its approach allows for a systematic evaluation of the overall performance of the entire system by isolating the many diverse functions of the subsystems or modules for closer scrutiny. It is now clear that the input/output comparison approach is applicable to a single module, a group of modules, or even to the entire system. This is demonstrated by the previous example of voice at the input transducer of the transmitter being compared with the output voice from the receiver output transducer.

The simulation of CNI systems can also be approached in this manner. The actual methods of simulation are diverse enough to allow total system simulation to be realized earlier than usually anticipated by mixing and matching hardware and software in the model. A possible

example is a transmitter that is completely simulated in modular fashion. Suppose a hardware design change is implemented in the first amplifier stage. With proper interfacing the new hardware module could be substituted for its simulation module in the system simulation sequence, and its effects on total system performance tested in real time or non-real time depending on the situation.

Different types of software simulations for individual modules can take the form of: (1) pure mathematical models, (2) convolution with calculated module impulse response functions, (3) convolution with actual measured module impulse response functions, (4) circuit analysis programs like SPICE and (5) actual hardware modules. Of course hybrid simulations including some or all of the above increase the flexibility of the modular approach. Keeping these thoughts in mind, the term system or module used in the rest of this report will imply the most general interpretation.

Some traditional methods of CNI system evaluations include measurements usually performed in susceptibility analysis. These will not be discussed in this report but they should be considered as an important part of any evaluation approach. The Interactive Communication Simulator (ICS)<sup>8</sup> can also be helpful in the evaluation process. The most advanced Network Analyzer now on the market seems to be the HP5810. This Network Analyzer is completely programmable and while it is costly it is so efficient that it will be difficult for any one group to give it 100% utilization. The possibility of time sharing seems reasonable. There are three on order from AFWAL at this time. These analyzers can be useful in characterizing non-linear systems also.

#### IV. DIGITAL SIGNAL PROCESSING

The usual approach to the analysis of linear time invariant systems<sup>3,4,5</sup> begins by defining some function of time  $x(t)$  to be an input signal to the system whose unit impulse response is given by  $h(\tau)$ . The output signal  $y(t)$  caused by the input  $x(t)$  is then given by the convolution integral

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t-\tau) d\tau. \quad (1)$$

In the frequency domain this relationship reduces to

$$Y(f) = H(f) X(f) \quad (2)$$

where

$$Y(f) = \int_{-\infty}^{\infty} y(t) e^{-j2\pi ft} df. \quad (3)$$

Note that equation (2) leads to the relationship

$$H(f) = \frac{Y(f)}{X(f)}. \quad (4)$$

Then, the Fourier Transform of the unit impulse response which is referred to as the frequency response function is simply the ratio of the Fourier Transform of associated input and output signals. Note that in general both amplitude and phase information is available. Equation (3) is the usual definition of a continuous Fourier Transform and the transformation equation is the same for  $X(f)$  and  $H(f)$ . However, in practice experimental approaches do not record continuous values of  $x(t)$  and  $y(t)$ , but record samples  $x(k)$  and  $y(k)$  at sample times  $\Delta t$  apart, and

for record lengths  $T = \Delta t \cdot NS$ , where  $T$  is the length (time) for a record of  $NS$  discrete samples. Then the Discrete Fourier Transform (DFT) is defined by<sup>9,10</sup>

$$Y(m) = \frac{1}{NS} \sum_{k=0}^{N-1} y(k) e^{-j \frac{2\pi}{NS} mk} \quad (5)$$

where all terms have their usual meaning. The Fast Fourier Transform (FFT) is simply an algorithm which computes the DFT quickly and efficiently by taking advantage of the mathematical properties of the sum in Equation (5).

The power spectral density function or autospectral density function or simply the auto spectrum of the signal  $y(t)$  is given by

$$S_{YY}(f) = Y^*(f) Y(f). \quad (6)$$

This function is the Fourier Transform of the auto-correlation function<sup>4,5</sup> of  $y(t)$ . However, in practice Equation (5) is really only an estimate  $Y_i(m)$  of  $Y(f)$ , gotten from the  $i^{\text{th}}$  record. A more reliable estimate<sup>4</sup> of  $S_{YY}(f)$  is realized by taking an average of many ( $NR$ ) records, where the discrete experimental values of  $S_{YY}(f)$  are given by

$$S_{YY}(m) = \lim_{NR \rightarrow \infty} \frac{1}{NR} \sum_{i=0}^{NR} Y_i^*(m) Y_i(m). \quad (7)$$

This relationship also holds for  $S_{XX}(m)$ . Another useful relationship called the cross spectral density function or simply the cross spectrum is defined by

$$S_{XY}(m) = \lim_{NR \rightarrow \infty} \frac{1}{NR} \sum_{i=0}^{NR} X_i^*(m) Y_i(m), \quad (8)$$

and  $S_{YX}(m) = S_{XY}^*(m)$ . The cross spectrum is the fundamental function relating the two signals to each other<sup>11</sup>. The cross spectrum is the Fourier Transform of the cross-correlation function<sup>4,5</sup> of  $y(t)$  and  $x(t)$ . Another valuable measure of the relationship between  $x(t)$  and  $y(t)$  is given by the coherence function

$$\gamma^2(f) = \frac{|S_{XY}(f)|^2}{S_{XX}(f) S_{YY}(f)} \quad (9)$$

The coherence function is related to the correlation coefficient function<sup>4,5</sup> but is not equal to its Fourier Transform. The coherence function can be interpreted to be a normalized cross spectral density function<sup>11</sup>. Similar to the correlation coefficient, the coherence  $\gamma^2(f)$  of the signals  $x(t)$  and  $y(t)$  is a function which on a scale from 0 to 1 measures the degree of linear relationship between the two signals at any given frequency  $f$ . For discrete data

$$\gamma^2(m) = \frac{|S_{XY}(m)|^2}{S_{XX}(m) S_{YY}(m)} \quad (10)$$

as might be expected. This relationship implies the desirability of a dual channel FFT Waveform Analyzer or system analyzer.

The averaging process of Equations (7) and (8) lead to some desirable relationships when the measurement system is effected by noise. If Equation (2) is multiplied by the complex conjugate of  $X(f)$

$$X^*(f) Y(f) = X^*(f) H(f) X(f) \quad (11)$$

here  $X^*(f)Y(f)=S_{XY}(f)$  and  $X^*(f)X(f)=S_{XX}(f)$ , it is seen that:

$$H(f) = \frac{S_{XY}(f)}{S_{XX}(f)} \equiv H_1(f). \quad (12)$$

Also if Equation (2) is multiplied by the complex conjugate of  $Y(f)$

$$Y^*(f) Y(f) = H(f) Y^*(f) X(f) \quad (13)$$

where  $Y^*(f)Y(f)=S_{YY}(f)$  and  $Y^*(f)X(f)=S_{YX}(f)$ , it is seen that:

$$H(f) = \frac{S_{YY}(f)}{S_{YX}(f)} \equiv H_2(f). \quad (14)$$

Furthermore if the absolute value squared is taken of both sides of Equation (2)

$$Y^*(f) Y(f) = H^*(f) X^*(f) H(f) X(f), \quad (15)$$

it can be seen that

$$|H(f)|^2 = \frac{S_{YY}(f)}{S_{XX}(f)} \equiv |H_a(f)|^2. \quad (16)$$

Note that the definition of  $H_1(f)$  and  $H_2(f)$  lead to the relationship,

$$\frac{H_1(f)}{H_2(f)} = \frac{S_{YY}(f) S_{YX}(f)}{S_{XX}(f) S_{YY}(f)} = \frac{|S_{XY}(f)|^2}{S_{XX}(f) S_{YY}(f)} = \gamma^2(f). \quad (17)$$

We therefore have several different methods of measuring the frequency response function of a system. If the system is perfectly linear and time invariant, then

$$H_1(f) = H_2(f) \quad (18)$$

and

$$|H_1(f)| = |H_2(f)| = |H_a(f)|. \quad (19)$$

If these equalities do not exist, the system is either not linear and time invariant or there is noise present.

Figure 2 shows a flow diagram for a typical dual channel FFT Waveform Analyzer or system analyzer. Figure 3 gives some examples of this approach in the presence of noise. In Figure 3  $m(t)$  is noise introduced at the input side of the system and  $n(t)$  is noise introduced at the output side of the system. As can be seen in the Figure, valuable system information is available even in the presence of noise. The ideal System Analyzer will measure either  $H_1(f)$ ,  $H_2(f)$ ,  $H_a(f)$  or a suitable combination of them.

The minimum requirement for digital signal analyzers to avoid aliasing<sup>8-13</sup> is a sampling rate at least twice the bandwidth of the signal. However for good reproduction of sampled signals, requirements of eight to twenty times the bandwidth of the signal are given<sup>14</sup>. Some of the fastest digitizers on the market will take 1024 samples at a maximum rate of 200 MHz. This rate will yield 20 samples per cycle of a 10 MHz sinusoid for a total of 51 cycles per record and will take 5.12  $\mu$ s per record. It doesn't seem possible to achieve better than five bit accuracy at this rate<sup>14</sup>. Lower sampling rates will improve the number of significant bits per word. Many dual channel FFT System Analyzers are available up to 500 kHz but the TEK 7612D seems to be the best overall choice to achieve reliable sampling up to 10 MHz.

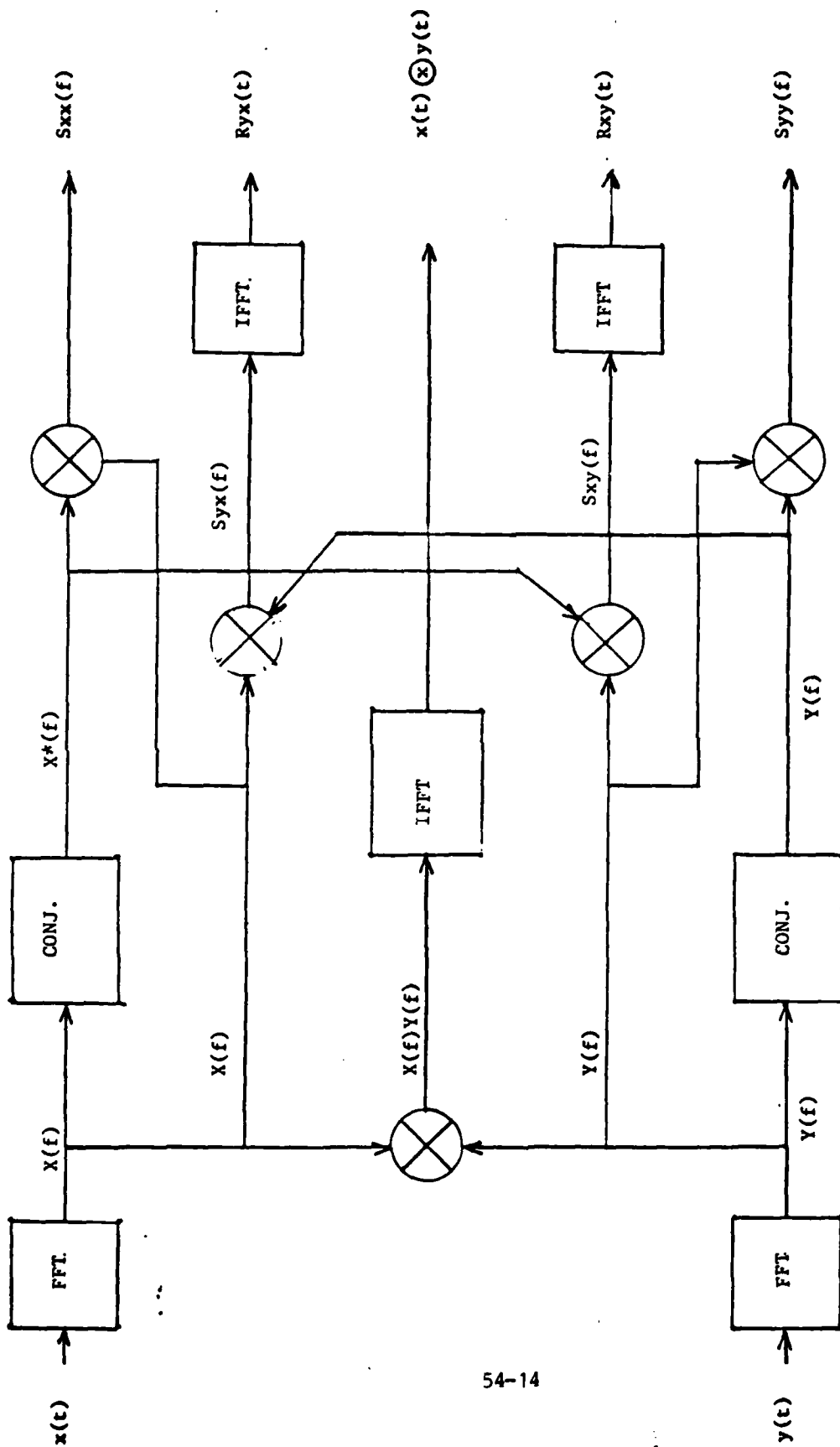


FIGURE 2. DUAL CHANNEL FFT SYSTEM ANALYSER



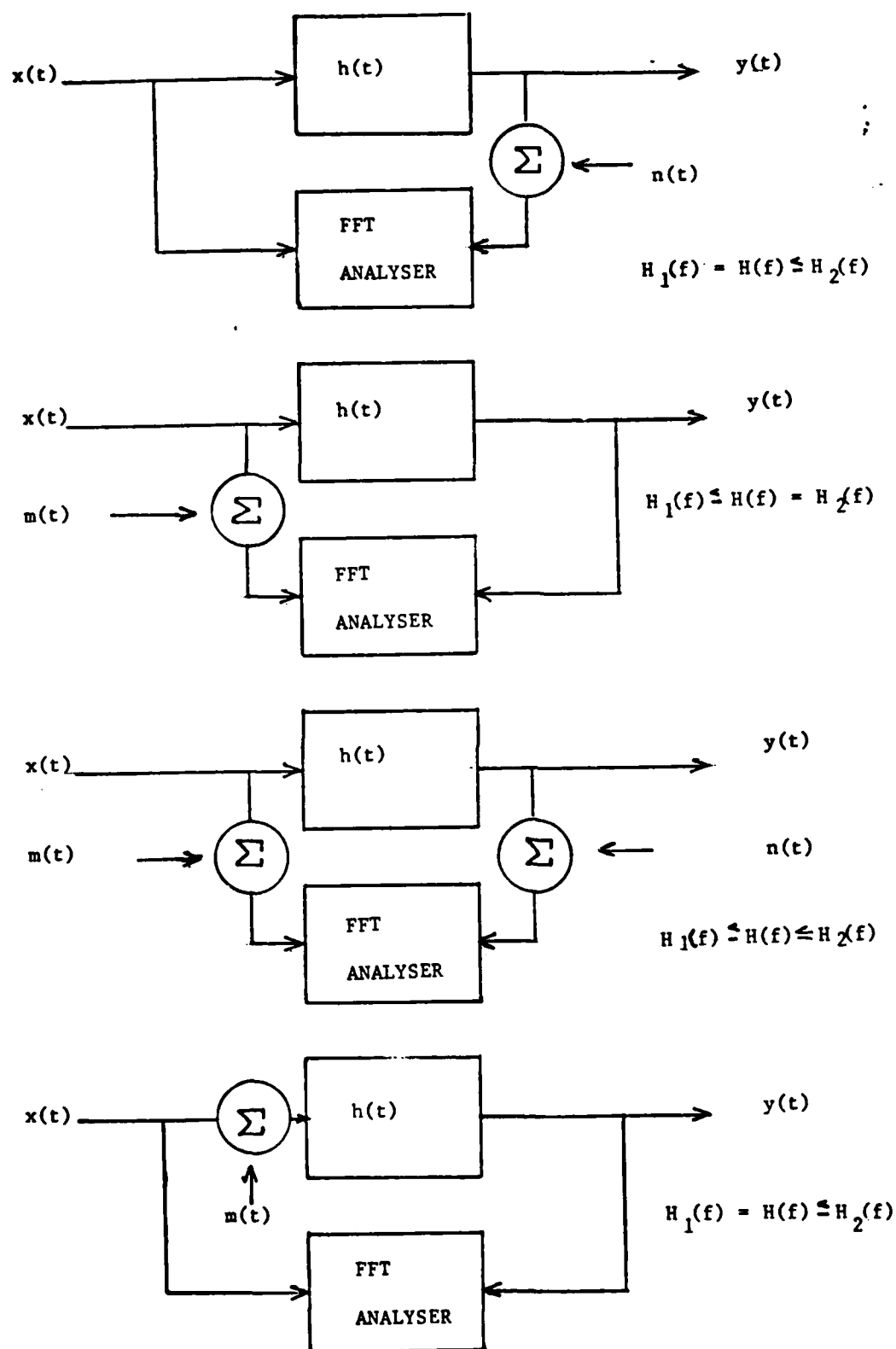


FIGURE 3. EFFECTS OF NOISE IN DUAL CHANNEL FFT SYSTEM ANALYSIS

#### FFT SYSTEM ANALYZER FOR AAAI

An attempt has been made during this effort to create an FFT system analysis capability with existing software and hardware presently available in AAAI. There were several FORTRAN programs written not all of which were completely debugged during this effort. Testing, debugging, expanding, and refining these programs will be a major part of the effort set forth in the proposal being prepared for the Research Initiation Program sponsored by AFSOR. The design of the FFT system analyzer facility using an AP120B Array Processor driven by the Harris computer is shown in Figure 4. The input to all the programs is read from the Harris system disk. The format of all functions stored on the disk is expected to be (3X, E15.7) for real variables and (2(3X, E15.7)) for complex variables. The first record of each file is the number of records that follow. The format for the first record is (3X, I10). All the programs in Figure 4 read their input functions from the disk expecting the format scheme as defined above. All the programs that create output functions, write these output functions on the disk using the same format scheme described above. Therefore, the output of any program may be used as the input for any other program. At least their formats are compatible. Proper logic is dictated by the user. The only programs that do not follow this scheme are the plotting routines which do follow the input scheme but create no output except a graph displayed on the TEK 4016-1 graphics terminal. A hard copy of the graph if desired is available through the TEK 4610.

The reason for the compatibility between program input and output formats is obvious. The reason for designing all inputs to be read from

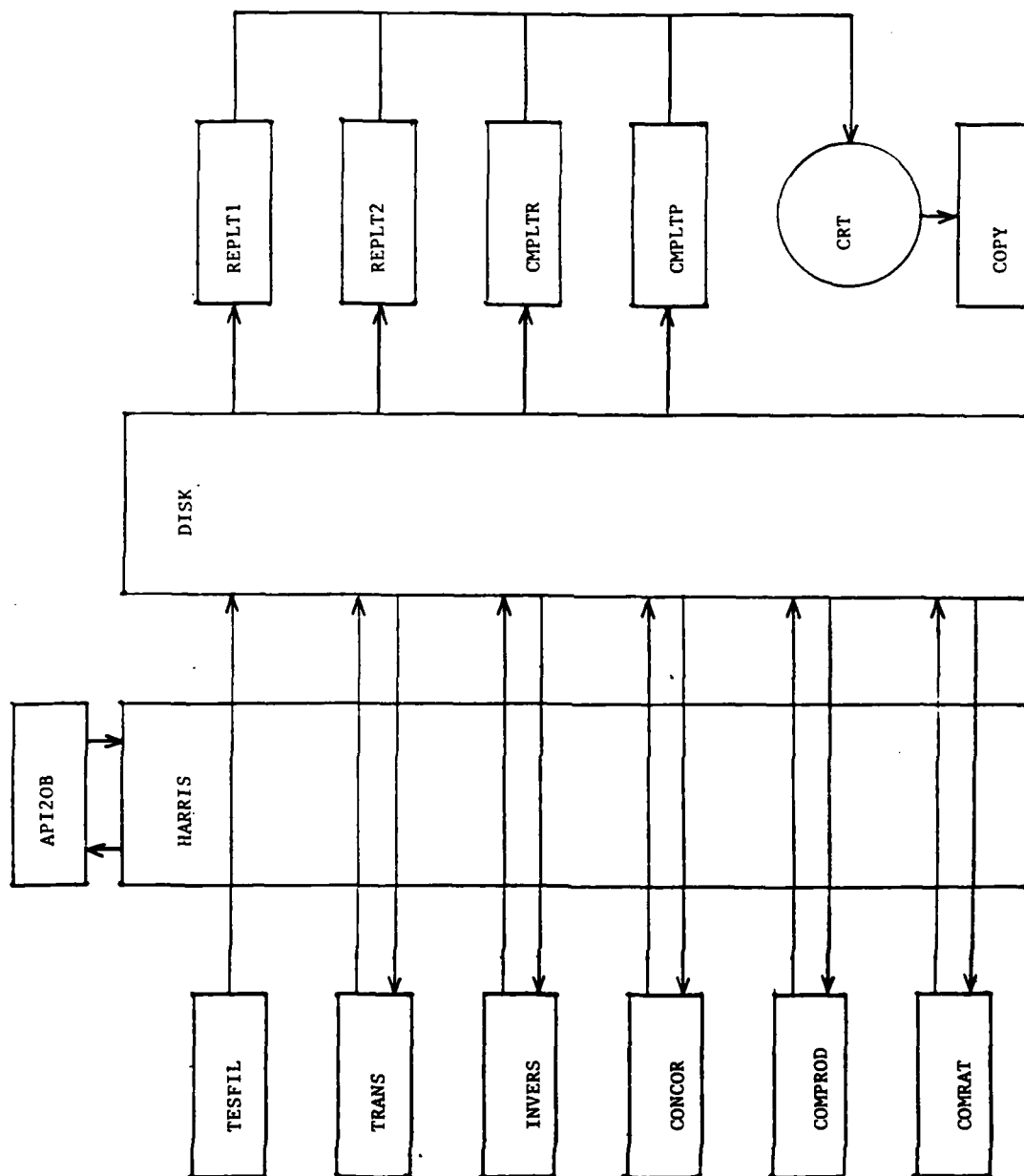


FIGURE 4. DUAL CHANNEL FFT SYSTEM ANALYSIS FACILITY IN AAAI-4

the disk was an attempt to allow the programs to accept experimentally generated data provided that the data can be written on the disk. Recommendations concerning methods of transferring real experimental signals to the Harris computer will follow.

A description of each program will be presented here. Complete documentation and FORTRAN statement listings are in the possession of Mrs. Lydia Harris of AAAI-4.

TESFIL: This program creates functions to be used as inputs for the other programs and writes them on the disk. The name of the new file, the number of points, the type of function and the amplitude are determined by the user. Choice of functions are: square wave, sinusoid, unit step, delta function and comb. They may all be written on the disk with either format, (3X, E15.7) denoting a real function or (2(3X,E15.7)) for a complex function with all imaginary parts equal to zero. Up to 1024 points may be generated for any function created by TESFIL.

REPLT1: This program will read a real function with format (3X, E15.7) from the disk and plot it on the Tektronix CRT. A hard copy may be requested from the keyboard of the CRT. The input for this program is simply the file name of the desired function on the disk.

REPLT2: This program is similar to REPLT1 but it will plot two real functions on the same graph for comparison. The graphs are automatically scaled to the largest value of the second file read in. The only inputs required are the names of two real functions stored on the disk. A listing of this program is included in the Appendix.

CMPLTP: This program reads a complex function from the disk assuming that its format is (2(3X, E15.7)) in the polar form (MAGNITUDE, PHASE). The program plots two curves on the same graph scaled to the largest magnitude. The magnitude of the phase is plotted assuming its units are radians modulo  $\pm \pi$ . The only input to CMPLTP is the name of the file to be plotted.

CMPLTR: This program reads a complex function from the disk assuming that its format is (2(3X, E15.7)) in the rectangular form (REAL PART, IMAGINARY PART). The program plots two curves on the same graph scaled to the largest value of the real part. The only input to CMPLTR is the name of the file to be plotted.

Further information about the plotting facility can be found in the file PLTMAN/T stored in the Harris computer.

These plotting facilities are invaluable in checking large input and output files, for debugging, and for interpreting results. They were written early on in the effort, as soon as it was discovered how much time was needed to read 1024 values of data. Figure 5 is an example of a graph of a gate function and Figure 6 is its auto correlation function.

TRANS: This program reads a time domain function (TDF) from the disk, takes the FFT according to Equation (5) and writes the resulting frequency domain function on the disk under a file name specified in the input. The user has a choice of the form of the output file. The frequency domain function (FDF) which is in general complex can be written to the disk either in rectangular form (REAL PART, IMAGINARY PART) or in polar form (MAGNITUDE, PHASE:MODULO $\pm \pi$ ). The inputs to this program are the name of the input file, the name of the output file and

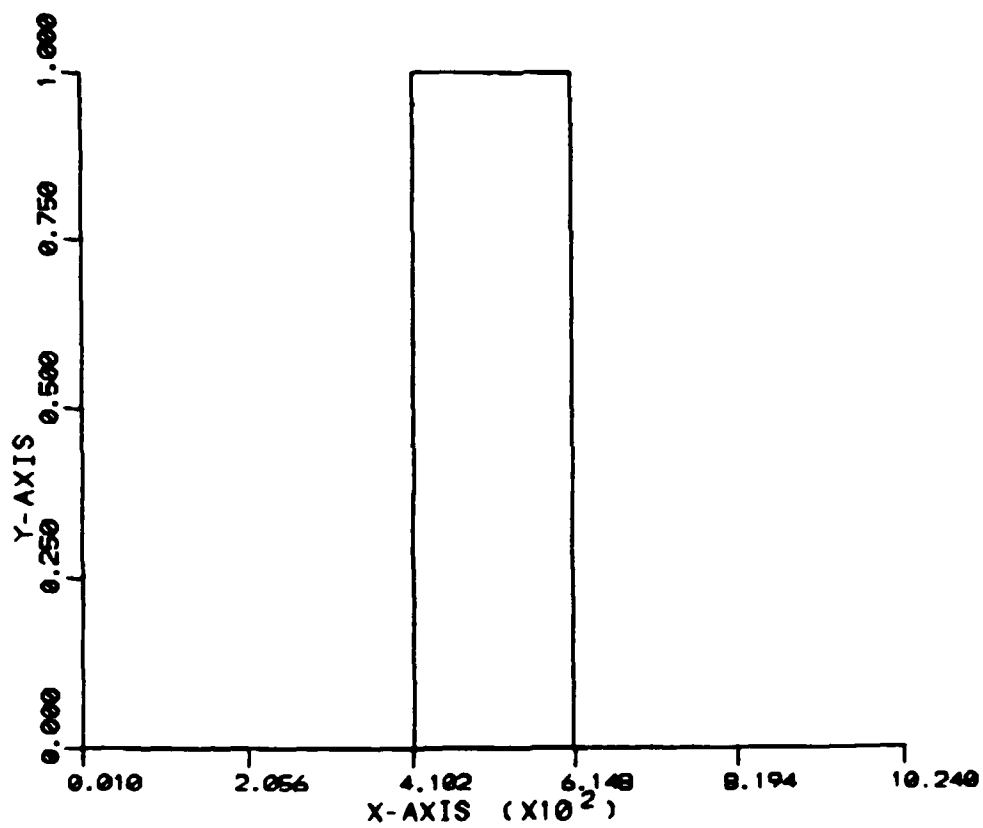


FIGURE 5. GRAPH OF A DIGITAL SIGNAL (GATE FUNCTION) HAVING 1024 SAMPLES.

the form of the output +1 denoting rectangular or polar form for the complex frequency domain function. A listing of this program is included in the Appendix.

INVERS: This program reads a complex frequency domain function from the disk and performs an Inverse Fast Fourier Transform (IFFT) according to the relation,

$$y(k) = \sum_{m=0}^{N-1} Y(m) e^{j \frac{2\pi}{NS} mk} \quad (20)$$

which is the inverse operation of Equation (5). The resulting time domain function is expected to be complex. INVERS as it exists at this time only writes the complex time domain function on the disk in the rectangular form. The input to this program is just the file name of the frequency domain function to be inverse transformed and the file name of the output time domain function.

In hindsight, it would have been more convenient for the user if TRANS and INVERS were combined into one program with input formats and output formats the same, i.e. (2(3X, E15.7)) with real functions having all imaginary parts equal to zero. The FFT and IFFT function could then be determined by an input flag like +1 designating forward transform or inverse transform.

CONCOR: This program takes two time domain functions from the disk and either does a correlation or a convolution operation on the two input functions, then writes the output function on the disk. An input parameter +1 determines the correlation option. If the number of samples in the input functions are NPTS1 and NPTS2 respectively, the

output convolution or correlation has  $NPTS1 + NPTS2 - 1$  samples. The inputs to this program are the file names of the two input functions and the file name of the output function. If the file names of the two inputs are the same, the output will be the autocorrelation or autoconvolution of the input. Figure 6 is a graph of the auto correlation of the gate function in Figure 5.

COMPROD: This program takes two complex frequency domain functions from the disk, calculates their product term by term, and writes the resulting complex frequency domain function on the disk. Remember that multiplication of two functions in the frequency domain is equivalent to convolving the functions in the time domain. However, care must be taken to resolve the anomaly that if this product is taken for two functions of length  $N$  in the frequency domain, the frequency domain product still has only  $N$  points. In the discussion of CONCOR one can see that this is not the case in the time domain. There are techniques to resolve this anomaly, but none have been implemented during this effort. A similar circumstance holds for the correlation. The inputs to COMPROD are the file names of the two complex FDF's to be multiplied and the output file name of the product. COMPROD will be useful for evaluating autospectra as in Equation (8), cross spectra as in Equation (9), and the denominator of Equation (10) for the coherence.

COMRAT: This program takes two complex FDF's from the disk and evaluates their complex ratio. Used in conjunction with TRANS and COMPROD this program is useful for evaluating  $H_1(f)$ ,  $H_2(f)$ ,  $H_a(f)$  and  $\gamma(f)$  from Equations (12), (14), (16), and (17) respectively. The inputs to COMRAT are the file names of two FDF's and the file name of the



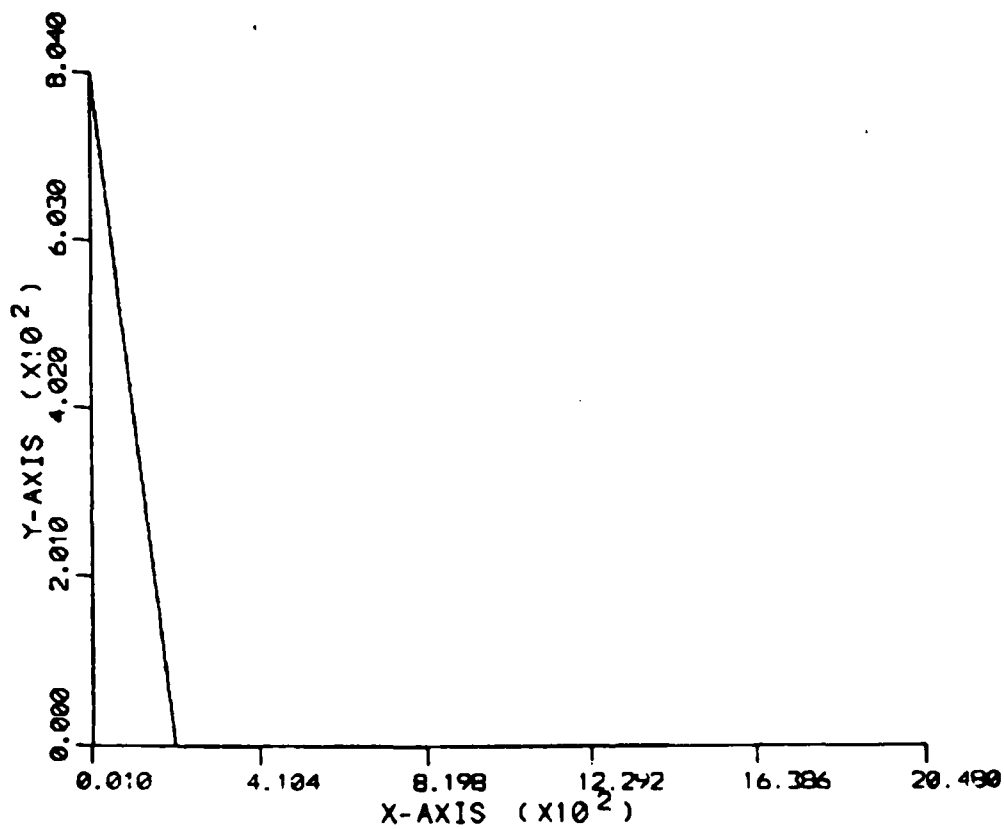


FIGURE 6. AUTO CORRELATION FUNCTION (2048 SAMPLES) OF DIGITAL SIGNAL IN FIGURE 5.

output ratio. The file whose name is input first, is the numerator of the ratio.

As can be seen from Figures 2 and 4, with some further effort the Harris and API20B plus the above mentioned software can function as a non-real time FFT system analyzer. An excellent article describing the FFT and the pitfalls encountered by the user was written by G.D. Bergland<sup>15</sup>. It is recommended reading for anyone interested in the programs discussed here. A more recent article<sup>16</sup> on applications of  $H_1(f)$  and  $H_2(f)$  can be found in the March 1984 issue of Sound and Vibration.

With the capability of reading files from the disk, the above digital signal processing facility can analyze either analytically generated signals or real measured signals, as long as the measured signals can be sent to the Harris disk. AAAI needs a facility to send experimentally digitized real signals to the Harris for processing. A digitizer similar to the TEK 7612D is recommended at this time. There is a "rumor" afield that HP will bring a competitive Programmable Digitizer to market soon and that the next generation of TEK 7612D will be available within a year.

An investigation was undertaken to find what on-site hardware was available to expedite the transfer of actual measured signals to the Harris disk. Figure 7 shows the available equipment and identifies the missing links. The TD1007J boards and the HP9836S computer are available in AAAI-4 at the present time. The Zenith H/Z100 computer is expected soon. Latches can be made in-house and a smart telephone modem could be used to transfer data from the HP9836S or Zenith H/Z100 to the

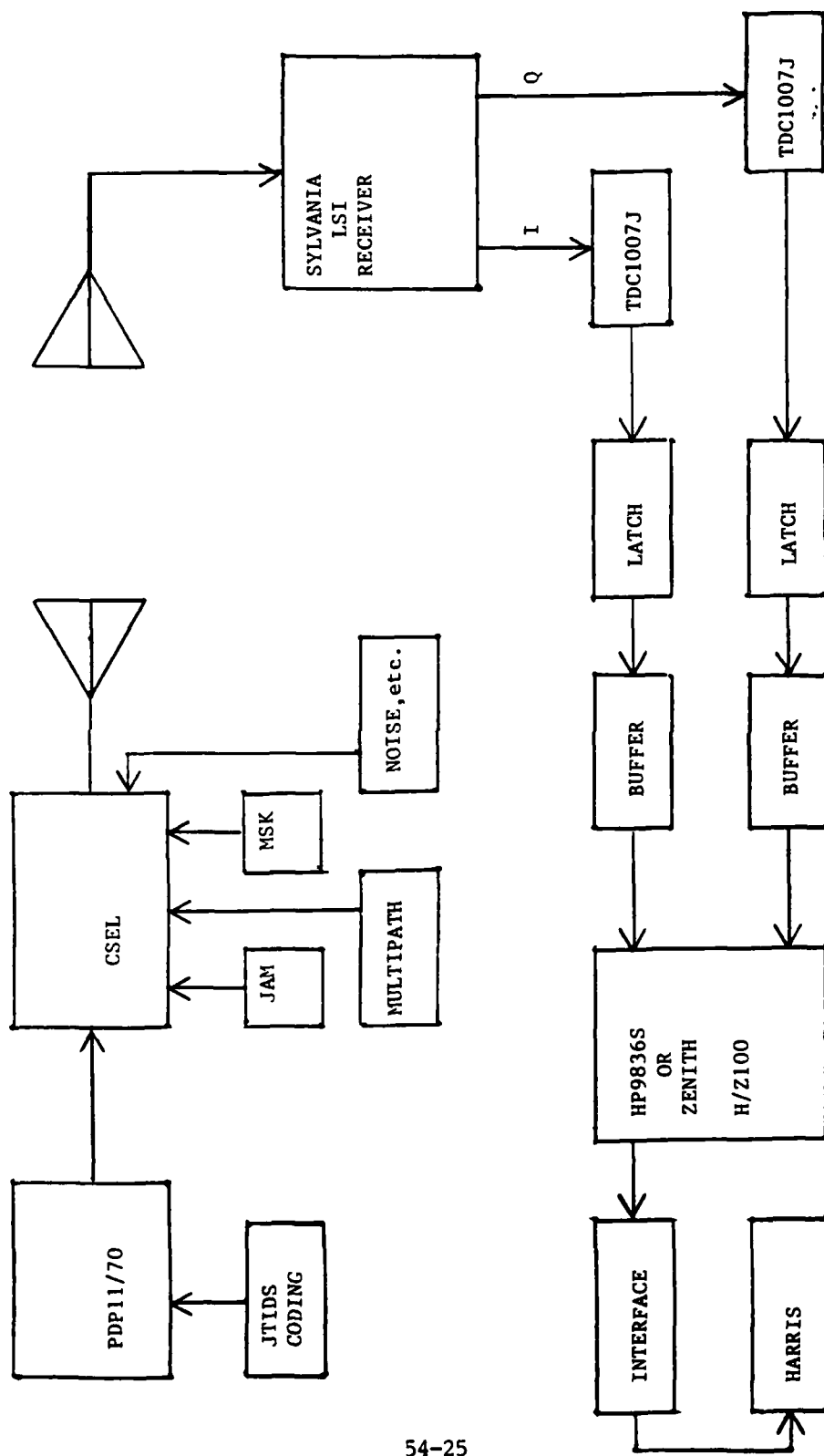


FIGURE 7. METHOD OF SENDING REAL RF SIGNALS TO HARRIS DISK.

Harris in non-real time. The rate buffers or memory buffers or a direct memory access (DMA) interface would still be needed. This route may not be practical.

#### V. THE MISSION OF AAAI-4

Figure 8 taken from a TRW report on the status of IESS lists the CNI services in ICNIA. It is recommended here that AAAI-4 choose a few critical services from this and concentrate their analysis and evaluation efforts on those specific services. This limitation of the scope of effort may make the approach more manageable. The services most likely to cause problems are identified by AAAI-4 to be the Joint Tactical Information Distribution System (JTIDS), the Global Positioning System (GPS) and one of the Identification Friend or Foe services (MARK XV IFF).

This selection is somewhat convenient since the Minimum Shift Keying (MSK) modulation facility being added to the Communication System Evaluation Laboratory (CSEL) in AAAI-4 will enable AAAI-4 with some enhancement of CSEL's hopping rate to generate JTIDS compatible signals. Furthermore, the GPS Evaluation (GPSE) is still in the possession of AAAI-4 and can be used to generate GPS signals plus jam.

At this time the MSK hardware for CSEL is being constructed by R. Clark and D. Welbaum. The immediate intention is the use of CSEL for sending MSK modulated signals to the Sylvania LSI Receiver. The output of the LSI Receiver will be baseband with both in phase (I) and in quadrature (Q). This signal will then have to be digitized and stored in computer memory. These signals can be generated with the many

BASISLINE CNI SIGNAL SOURCES	ICNIA SIGNAL (GENERATED)	NUMBER OF SIGNALS REQUIRED	NUMBER OF SIGNALS BY AIM	UNIT (HMM)	COMMENTS
JTIDS	YES	2	2	L	ICNIA AIM provides functions of two (2) JTIDS terminals. Software and includes interleaved processing and scheduling of pulsed waveforms. Mode also supports simulation of multiple range delays via rescheduling of pulses
EJS	YES	1	1	Low	ICNIA will provide functions of future Air Force AJ UHF communications
PLRS	YES	1	1	Low	Software modified in ICNIA AIM to emulate Master Unit's control of PLRS user units and synchronization of PLRS net.
HAVE QUICK	YES	1	1	Low	ICNIA provides functions of HAVE QUICK
SINGARS	YES	1	1	Low	ICNIA provides functions of SINGARS
GPS	NO	4	0	L	Modification of ICNIA AIM to meet simulation requirements is not practical o Much new hardware would be required to generate 8 GPS signals o Without SWS, extensive software would have to be redeveloped
TACAN	NO	1	0	L	TACAN simulation requirements are best achieved using off-the-shelf, supportable, and cost effective test equipment
MARK XII IPP	YES	1	1	L	ICNIA provides functions of MARK XII IPP
ILS/VOR	NO	1	0	Low	ILS/VOR simulation requirements are best achieved using off-the-shelf, supportable, cost effective test equipment.
TCAS	YES	1	1	L	ICNIA provides functions of TCAS
MARK XV IPP	YES	1	1	L	ICNIA provides functions of MARK XV IPP
APRATCOM	YES	1	1	Low	Software modification in ICNIA AIM escape up-link and down-link modulation formats to allow back-to-back APRATCOM terminal communication without relay satellite.
HP, VHF, and UHF	YES	3	0	Low	ICNIA provides functions of HP, VHF, and UHF voice communications.
ACMI	NO	1	0	L	ACMI simulation requirements are best achieved using off-the-shelf, supportable, cost effective test equipment.

FIGURE 8. A LISTING OF THE CNI SERVICES CONTAINED IN ICNIA

atmospheric effects now available in CSEL and can be used to test the JTIDS software simulation model of AAAT-2 in non-real time, for the present. Eventually CSEL can be used to test JTIDS receivers in real time. Again the weakest link here is the digitizer. Since the baseband MSK spread JTIDS signal is expected to have a bandwidth less than 8 MHz, a digitizer similar to the TEK 7612D would be adequate.

One of the disadvantages of digital signal processing approaches is the inability to digitize signals with components above 10 MHz or so. Therefore, all digital simulation of CNI systems must ignore the high frequency signal paths or the RF. The RF is where all atmospheric effects occur. Furthermore, the RF path is where interference, jam, multipath, fade, doppler, and other real transmission problems occur. It is obvious therefore that all digital CNI simulations must have an analog RF module to bring them into the real world. Fortunately CSEL could be a convenient solution to this problem.

It seems apparent that IESS will not have a jamming capability for testing ICNIA. Here again both CSEL with the University of Michigan Jammer and GPSE with its jammers could be used to support IESS with jamming capabilities.

Figure 9 gives a possible design for a real time active jamming facility to support IESS in its testing scenarios for ICNIA. This recommendation is based on the concept of the former Sweep Lock Jammer Simulator (SLJS) which was built around the Programmable Signal Processor (PSP).

The conclusion implied here seems to indicate AAAI-4 can support the IESS and ICNIA analysis and evaluation effort with the digital signal

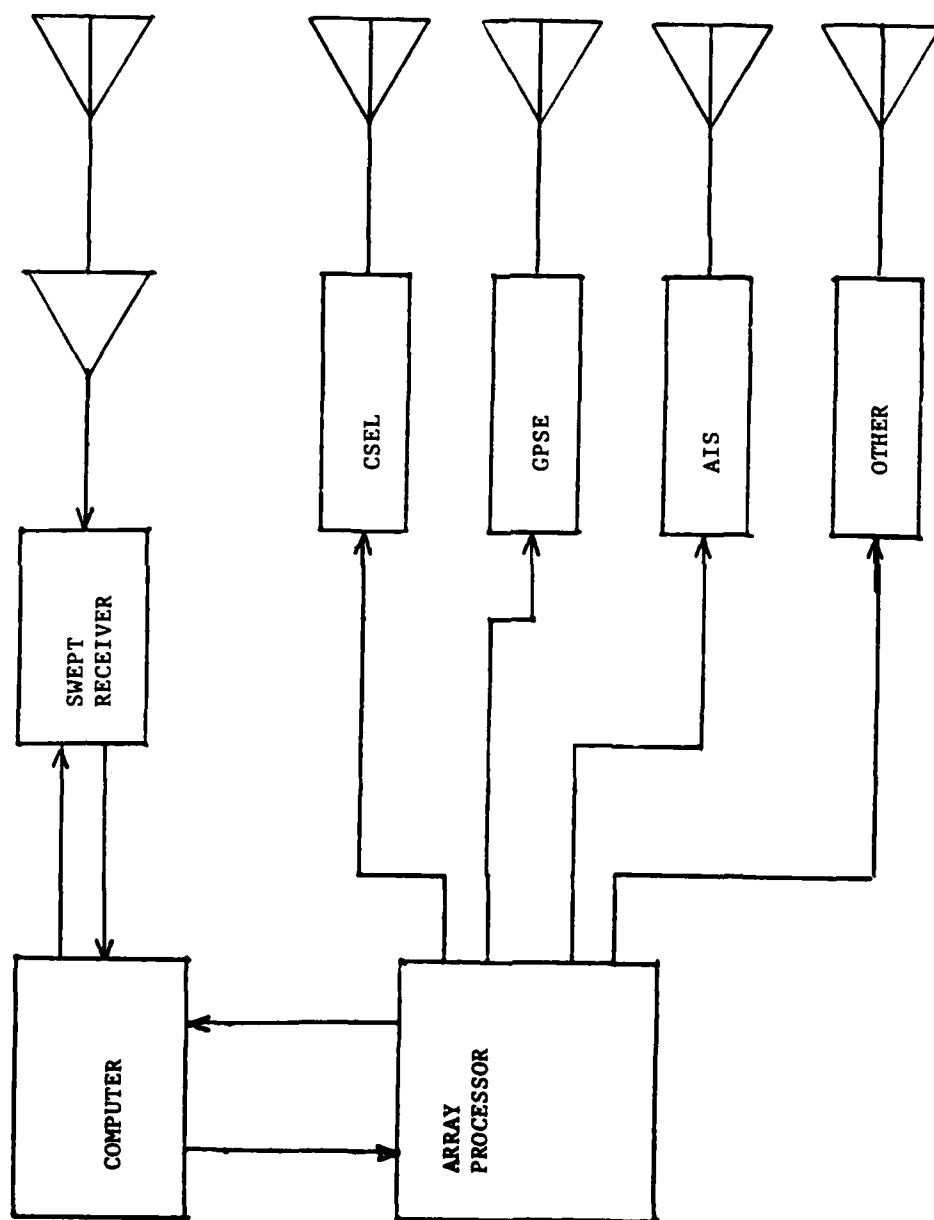


FIGURE 9. A PROPOSED ACTIVE REAL TIME JAMMING FACILITY USING IN HOUSE HARDWARE.

processing approach of the FFT systems analysis capability in the Harris and AP120B, provided a suitable digitizer is acquired. Also, AAAI-4 can adequately support the IESS and ICNIA analysis and evaluation effort by supplying the RF path with most environmental effects needed to test CNI systems. With the addition of the design in Figure 9, AAAI-4 could support an important function of IESS for which there seems to be no other provision, that is the concept of the active real time programmable jammer.

#### VI. RECOMMENDATIONS

The number of recommendations included in this section should not be construed as a criticism of AAAI, but rather they are a result of many informative discussions with the group. Further, many of these recommendations cannot be considered as original but simply a re-statement of suggestions and plans made by AAAI while the author was on-site. Some of these suggestions will be expanded in greater detail in the proposal being prepared for the AFSOR Research Initiation Program. This proposal is being written with the support of the AAAI-4 staff.

1. The existing digital signal programs in the Harris computer should be completely tested, debugged, and refined to suit the needs of AAAI-4. These programs should be expanded and combined so that the Harris AP120B could function similarly to a dual channel FFT Wave Analyzer.



2. This process would be expedited and the efficiency and effectiveness of the above effort greatly increased by the acquisition of the signal analysis package for the AP120B.

3. To improve the existing algorithms for the AP120B and to write new algorithms, the FORTRAN package would be valuable. These packages would be versatile enough to be used on the next generation of Floating Point array processors which are already on the market.

4. A high frequency dual channel digitizer like the TEK 7612 should be acquired to digitize baseband signals with adequate sampling rates and store them in a memory that can be conveniently read by the Harris computer at reasonable transfer rates.

5. An interim method of sending digitized signals may be attempted using the HP9836S computer already available in AAAI-4; however the interface from the A/D boards to the HP9836S and a suitable interface between the HP9836S and the Harris computer are not presently available on-site. Another potential interim solution could be the H/Z-100 computers which will be available on-site in the near future. Again interfacing is a problem that must be solved. In both cases a slow transfer rate from the computers mentioned above and the Harris could be accomplished by telephone modem.

6. Any existing simulation models of the IESS or ICNIA systems or their subsystems should be acquired to study new and existing methods of analysis and evaluation.

7. A few subsystems of the IESS and ICNIA system which involve new technology or seem most likely to pose problems should be identified. The AAAI-4 effort should be concentrated on these systems to optimize

the evaluation process. JTIDS, GPS and MARK XV IFF seem to be likely candidates. Also, the Universal Matched Filter of TRW and the Programmable Transversal Filter of ITT & TI along with the Reflective Array Compressor for the Fourier Domain Processor of CSDL seem to be likely candidates in the hardware evaluation area.

8. As soon as any hardware modules for the IESS and ICNIA subsystems are constructed they should be acquired for immediate study, so that positive feedback can be given the contractors during the development process.

9. A network analysis system similar to the HP8510A should be acquired or time shared for the analysis of IESS and ICNIA hardware as it becomes available.

10. The CSEL system should be enhanced especially by adding MSK modulation that is compatible with JTIDS. The CSEL and GPSE systems may be useful in supplementing IESS particularly with JAM capabilities. For real time response scenarios they would have to be driven by an array processor and a computer. This computer could be the existing PDP 11/70 provided the present memory is expanded.

11. To insure continuity and uniformity in the ongoing analysis and evaluation of systems and subsystems, some test signals, reference signals, and their expected or required responses under various environmental conditions should be archived for future reference. These signals could be created in-house and/or collected from contractors. To better simulate environmental conditions, signals could be recorded in the field.

12. Through the Follow-On Grant Program the continued services of the author may be utilized to test, debug, refine, and expand the software created for the Harris AP120B digital signal processing facilities. Also training sessions for AAAI staff in using this facility for each of their special diverse applications could be provided under this program.

#### ACKNOWLEDGEMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and productive summer at the Air Force Wright Aeronautical Laboratory, Wright-Patterson Air Force Base, Ohio. He would like to thank the laboratory, in particular the Information Transmission Branch, for its hospitality and excellent working conditions.

The author also wishes to thank all those individuals who took the time and effort to supply him with information or documents concerning the IESS, ICNIA (TRW), ICNIA (ITT-TI), FDP (DRAPER), and AIS programs in particular Capt. E. Wallace, R.L. Harris, L.R. Nawman, Capt. C. Batchelor, and Capt. R. Colvin. Information about JTIDS was received from L. Gutman and J.W. Mayhan. Invaluable assistance with the Harris computer system was generously supplied by D. Harper, C. Hicks, J.C. Caffrey, 1Lt. A.M. Staufer, and M.E. Sturgel. Information and much useful advice about state of the art digitizers was contributed by C. Lynn and others. Dr. R.A. Phelps was very helpful in supplying answers to many detailed questions concerning usage and applications of the API20B. The immediate availability of the expertise of these parties greatly reduced the time needed to complete many aspects of this effort.

In AAAI-4 assistance in attempting A/D conversion came from Lt. R. Rovinsky and Msgt. R. Carpenter. Timely construction of an MSK modulation circuit for CSEL is being done by R.C. Clark and D. Welbaum.

Many productive hours at the Harris computer console were contributed by Mrs. L. Harris. Leadership for, and the coordination of all these diverse activities was ably managed by Mrs. D.E. Summers, without whose guidance and cooperation this program could not have been so successful.

#### REFERENCES

1. ICNIA Technical Proposal in Response to: F33615-83-C-1001.
2. IESS Technical Proposal in Response to: F33615-83-R-1109.
3. Aseltine, J.A., "Transform Method in Linear System Analysis", McGraw-Hill Book Co., Inc., NY, 1958.
4. Bendat, J.S., and Piersol, A.G., "Random Data: Analysis and Measurement Procedures", Wiley-Interscience, NY, 1971.
5. Bendat, J.S., and Piersol, A.G., "Engineering Applications of Correlation and Spectral Analysis", Wiley-Interscience, NY, 1980.
6. Williams, L., Wunsch, R.M., Zuraw, J.F., and Griesacker, P.B., "The Effect of Spatial Frequency Filtering on Optical Image Formation", Proc. Pa. Acad. Sci. 49 (1975) 177-182.
7. Griesacker, P.B., "Calibration of Wideband Optical Signal Processor", Final Report from SCEEE Summer Faculty Research Program to the AFOSR, 1981.
8. Modestino, J.W., "Interactive Simulation of Digital Communication Systems" IEEE, J.S.A.C., Vol. SAC-22, No. 1, January (1984) 51-75.
9. Oppenheim, A.V. and Schaffer, R.W., "Digital Signal Processing", Prentice-Hall, Englewood Cliffs, NJ, 1975.
10. Rabiner, L.R. and Gold, B., "Theory and Application of Digital Signal Processing", Prentice-Hall, Inc., NJ, 1975.
11. Biering, H. and Pedersen, O.Z., "System Analysis and Timedelay Spectrometry" Parts I & II, A Bruel & Kjaer Publication, Marlborough, MA, 1983.
12. Papoulis, A., "Circuits and Systems, A Modern Approach", Holt, Rinehart and Winston, Inc., NY, 1980.
13. Lathi, B.P., "Modern Digital and Analog Communications Systems", CBS College Publishing, NY, 1983.
14. Priv. Comm., Mr. Claude Lynn.
15. Bergland, G.D., "A Guided Tour of the Fast Fourier Transform", IEEE Spectrum, July 1969.
16. Upton, R., "Innovative Functions for Two-Channel FFT Analyzers", Sound and Vibration, Instrumentation Reference Issue, March (1984) 18-25.

A P P E N D I X

```

*PR ALL
*CAU---
C....PROGRAM **REPLY2**
      INTEGER*3 IXTIT(2),IYTIT(2)
      REAL A(1050),B(1050)
      DIMENSION X(1050),Y1(1050),Y2(1050)
      DIMENSION IHNM(11)
      DATA IXTIT / 'X-A','XIS' /
      DATA IYTIT / 'Y-A','XIS' /
      LUN=3
      CALL ASSIGN(6,3,0,IERR)
      PRINT, ' ENTER FIRST FILE NAME '
      READ(0,100) IHNM
100  FORMAT(11A3)
      IHNM(11)=0
      CALL ASSIGN(13,IHNM,IERR)
      IF(IERR.EQ.0) GO TO 10
      PRINT, ' NO DISK AREA WITH THAT NAME (10)'
      CALL EXIT
10  READ(13,200) NCT
200  FORMAT(3X,I10)
      DO 20 J=1,NCT
          X(J)=J-1
20  READ(13,201) A(J)
201  FORMAT(3X,E15.7)
      CLOSE 13
      DO 30 J=1,NCT
          Y1(J)=A(J)
30  PRINT, ' ENTER NAME OF SECOND FILE '
      READ(0,100) IHNM
      IHNM(11)=0
      CALL ASSIGN(13,IHNM,IERR)
      IF(IERR.EQ.0) GO TO 40
      PRINT, ' NO DISK AREA WITH THAT NAME (40)'
      CALL EXIT
40  READ(13,200) NCT
      DO 50 J=1,NCT
          READ(13,205) B(J)
205  FORMAT(3X,E15.7)
50  Y2(J)=B(J)
      CLOSE 13
      CALL PLOTS(I,LUN)
C SCALE FOR 5 INCHES ON DRAWING
      SX=5
      SY=4
      CALL SCALE(X,NCT,SX,XMIN,DX)
      CALL SCALE(Y2,NCT,SY,YMIN,DY)
      CALL PLOTNP
      XYSCL=1.0
      XC=1.0
      YC=1.0

```



C .OFFSET THE DRAWING UP ONE INCH AND OVER ONE INCH WITH SCALE 1:1

CALL OFFSET(XC,XYSCL,YC,XYSCL)

NXTIT=-6

THETA=0

CALL AXIS(0.0,0.0,IXTIT,NXTIT,SX,THETA,XMIN,DX)

NYTIT=6

THETA=90.

CALL AXIS(0.0,0.0,IYTIT,NYTIT,SY,THETA,YMIN,DY)

INC=1

M=0

ISYM=0

CALL LINE(X,Y1,NCT,INC,XMIN,DX,YMIN,DY,M,ISYM)

ISYM=1

CALL LINE(X,Y2,NCT,INC,XMIN,DX,YMIN,DY,M,ISYM)

CALL EXIT

END

SVU REPLT2/X

LIB 00000ALE\*PLT LIS \*LIBERY

BEGIN

```

      SGE TRANS/D
      S:IF !L1
      !L1 SGE TRANS/D
      SMO ECHO=ON
      SFREE ALL
      SS*SAUF77

      C...PROGRAM **TRANS**
      C...16 JULY 1984

      C...TDF(I)= TIME DOMAIN FUNCTION
      C...FDF(I)= FREQUENCY DOMAIN FUNCTION
      C.....NI= NUMBER OF SAMPLES IN TIME DOMAIN
      C.....T= TIME LENGTH OF ENTIRE MEASUREMENT
      C.....DELT= TIME BETWEEN SAMPLES/SAMPLE RATE
      C.....BW= BANDWIDTH OF SPECTRUM=1/DELT
      C.....DELF= DISTANCE BETWEEN SPECTRAL VALUES
      C...IAPLOC= LOCATION OF ORIGIN OF TDF IN AP120B
      C.RECT/POL= SPECTRA FORMAT: RECTANGULAR(1) OR POLAR(2)

      INTEGER NI,I,IAPLOC,IFMT,IMIN,IMAX,IAPNUM,IACON,ISTAT,
      I,IOEL,IAP1,IAP2,IAPM2,IAPT,IAPM,NIM,RECPOL,IERR,NPTS

      REAL TDF(1024),MAG(1024),PHAZ(1024),T,DELT,BW,DELF
      COMPLEX FDF(1024),FDFP(1024)

      DIMENSION IHNM(11)

      WRITE(3,*)'ALL REAL INPUTS ARE(E15.7)'
      IERR=0

      WRITE(3,*)'ALL INTEGER INPUTS ARE (I5)'
      WRITE(3,*)
      WRITE(3,*)'NO OF SAMPLES: NPTS(I5)= '
      READ(3,10)NPTS

      WRITE(3,*)'FORMAT OF COMPLEX SPECTRA: RECPOL(INTEGER 1 OR 2)= '
      READ(3,10)RECPOL

      WRITE(3,*)'TOTAL TIME OF MEASUREMENT: T(SECONDS)= '
      READ(3,20)T

      WRITE(3,*)'TIME BETWEEN SAMPLES/SAMPLE RATE: DELT(SECONDS)= '
      READ(3,20)DELT
      DO 999 J=1,NPTS
      MAG(J)=0.0
      PHAZ(J)=0.0

```

```

999 CONTINUE
10  FORMAT(15)
20  FORMAT(E15.7)
    CALL ASIGNI(6,3,0,IERR)
    PRINT,"ENTER INPUT FILENAME"
    READ(0,100)IHNM

100  FORMAT(11A3)
    IHNM(11)=0
    CALL ASSIGN(13,IHNM,IERR)
    IF(IERR.EQ.0)GO TO 39
    CALL EXIT
IN  39  READ(13,101)NPTS
101  FORMAT(3X,I10)
    WRITE(3,101)NPTS
    DO 110 I=1,NPTS
    READ(13,102) TDF(I)
110  CONTINUE
102  FORMAT(3X,E15.7)
    CLOSE 13
    PRINT,"ENTER OUTPUT FILENAME"
    READ(0,100)IHNM
    CALL CREATE(IHNM,0,0,0,IERR)
    IF(IERR.EQ.0)GO TO 40
    PRINT,"FILE ALREADY EXISTS ON DISK"
    CALL EXIT
    DU 40  CALL ASSIGN(13,IHNM,IERR)
    WRITE(13,101)NPTS
103  FORMAT(2(3X,E15.7))
    IAPLOC=0
    JPTS=2*NPTS
    IFMT=2
    IAPI=IAPLOC+NPTS+1
    IAP2=IAPI+1
    IAPM=IAP2+NPTS
    IAPT=IAP2+NPTS+1
    DU  IAPNUM=1
    IACTION=0
    IAPM2=IAPT+NPTS+1
    NIH=IFIX(NPTS/2+0.1)

    CALL APINIT(1,0,ISTAT)
    CALL APPUT(TDF,IAPLOC,NPTS,IFMT)
    CALL APWD
    CALL RFFTB(IAPLOC,IAPI,NPTS,1)
    CALL RFFTSC(IAPI,NPTS,3,1)
    CALL VATN2(IAPI,2,IAP2,2,IAPT,2,NPTS)
    CALL CVMAGS(IAPI,2,IAPM2,2,NPTS)
    DU  CALL VSORT(IAPM2,2,IAPM,2,NPTS)
    CALL APWR

```

```

CALL APGET(FDF,IAP1,NPTS,IFMT)
CALL APGET(FDFP,IAPM,NPTS,IFMT)
CALL APWD
GO TO (601,610)RECPOL
60  FORMAT(5X,I10,(4(3X,E15.7)))
601  DO 600 I=1,NPTS
      WRITE(13,103)FDF(I)
      WRITE(3,60)I,FDF(I),FDFP(I)
600  CONTINUE
621  FORMAT(9I8)
  NH  GO TO 620
610  DO 620 I=1,NPTS
      WRITE(3,60)I,FDFP(I),FDF(I)
      WRITE(13,103)FDFP(I)
900  FORMAT(3X,I10,2(3X,E15.7))
620  CONTINUE
      CLOSE 13
      WRITE(3,621)NPTS,JPTS,NIH,IAPLOC,IAP1,IAP2,IAPM,IAPT,RECPOL
      WRITE(3,10)IERR
      CALL APRLSE
      STOP

```

18

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PRODUCTION RATE VARIATIONS COST MODELS

Prepared by: Thomas R. Gullledge, Jr.

Academic Rank: Assistant Professor

Department and University: Department of Quantitative Business Analysis  
Louisiana State University

Research Location: Business Research Management Center  
AFBRMC/RDCB

USAF Research: Major James Weber

Date: August 14, 1984

Contact No: F49620-82-C-0035

PRODUCTION RATE VARIATIONS COST MODELS

by

Thomas R. Gullledge, Jr.

ABSTRACT

This research examines a model that may be used to estimate the cost impact of production rate changes on a timely and real world basis. The first part of the research analyzes current models that are being implemented by Air Force Systems Command, and it is demonstrated that these models are deficient in their theoretical definition and empirical construction. The second part of the research presents the current status of the development of a more realistic rate variations model for the F-16 aircraft program.

## I. INTRODUCTION

There has been much interest in recent years in the relationships among learning, production rate, and program costs. These relationships are of particular interest in the military acquisition of made-to-order equipment. At the outset of a weapon system program, a tentative monthly production schedule for the life of the program is negotiated between the contracting parties. This planning schedule covers the life of program, but formal contractual agreements between the Department of Defense and manufacturers usually cover only annual delivery requirements. Since annual funding allocations are characterized by political uncertainties, there is often a need to deviate from the planned production rate during the production phase of the program. Coincident with these rate changes, new cost estimates are required to support contract negotiations and additional funding requests.

There are many proposed methodologies for assessing the cost impact of a production rate change. A recent group of models constructed for Air Force Systems command is based on application of the Alchian<sup>1,2</sup> cost function. Even within this framework there is very little agreement about the relationships among learning, production rate, and program cost. While some studies, for example, Womer<sup>14</sup> and Womer and Gulledge<sup>15</sup> make assumptions concerning the cost impact of the above factors in developing models of optimal contractor behavior, others [e.g., Smith<sup>13</sup>, Large, et. al.<sup>12</sup>, Bemis<sup>4,5</sup>, Cox and Gansler<sup>8</sup>, Crouch<sup>9</sup>, Cox, et. al.<sup>7</sup>, Bohn and Kratz<sup>6</sup> address the problem directly by attempting to statistically estimate the above influences. In these latter studies contractor behavior is not a part of the modeling effort. The purpose of this research is to interface the two types of studies and show how estimates of the influences of learning

and production rate, which are often statistically unreliable, may be inconsistent with optimal contractor behavior. In addition, an alternative method for estimating rate variations is proposed.

## II. OBJECTIVES

The specific objectives related to the more general introduction are as follows:

1. to review and criticize the current Air Force Systems Command production rate model,
2. to propose and begin testing of a new production rate model that is more sensitive to the needs of Aeronautical Systems Division (ASD/ACCR).

It was understood from the beginning that it would be impossible to complete this project. That is, the present Air Force Systems Command model was developed by a team of researchers over a one year period with a \$150,000 budget. It would be impossible to review the existing system and devise a new system in a ten week research period. The agreed objective was to extend the work as far as possible.

## III. ESTIMATING COST IMPACTS

The learning curve, first formulated by Wright<sup>16</sup>, is an empirically specified relationship that yields declining units costs with increases in cumulative output. In recent years the more commonly used terminology has been "improvement curve." The improvement curve allows for reductions in cost that are due to factors other than repetition (learning). Gold<sup>10</sup> includes changes in product design, product mix, technology, facilities, etc. in this listing of other factors. Both the learning and improvement curves are described mathematically as



$$Z = \beta_0 X_1^{\beta_1} \quad (1)$$

where

$Z$  = the unit cost of the  $X_1$ th unit,

$\beta_0$  = a constant, commonly called the first unit cost,

$\beta_1$  = a parameter describing the slope of the quantity/cost curve,

$X_1$  = cumulative quantity produced.

Studies attempting to ascertain the relationships among production rate, learning, and program costs generally use the following augmented model:

$$Z = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \quad (2)$$

where

$X_2$  = some measure (usually a proxy) for production rate,

$\beta_2$  = a parameter describing the slope of the rate/cost curve.

Some researchers (e.g., Bohn and Kratz<sup>6</sup> call equation (2) the "rate analysis curve model."

The parameters in equation (2) are estimated from the log-linear form of the relationship using the linear regression or directly from (2) using nonlinear regression. Unfortunately both of these techniques often are plagued with statistical problems due to the collinearity between the independent variables,  $X_1$  and  $X_2$ . The source of this collinearity may be reasoned as follows. Often made-to-order production programs are characterized by initial production at a low rate with a gradual buildup in production rate throughout the program. In fact, given a learning curve, if the resource use rate does not decline, production rate must increase during the program. As a result cumulative quantity is highly correlated with production rate.

There are additional problems with the formulation described by equation (2). Cox and Gansler<sup>8</sup> and Bohn and Kratz<sup>6</sup> use lot size as a proxy

for production rate. However, the time required to produce a lot often changes over the life of the program. This is true in much of the data that these authors have analyzed, namely the C141 airframe program, the F102 airframe program, the F4 airframe program, and the Army's Black Hawk helicopter program. For example, lot sizes of 15 and 20 are not good proxies for production rate is the time horizons for the two lots are 1 and 4/3 years respectively.

Still, the main problem encountered is collinearity between output rate and cumulative output. Large, et. al.<sup>12</sup> concluded that the influence of production rate could not be estimated with confidence. Many later studies have also been unable to significantly measure the influence of production rate. Both positive and negative estimates of  $\beta_2$ , the slope of the rate/cost curve, have been obtained. Assuming that an increase in rate requires an increase in resources, a positive slope for the rate/cost curve implies decreasing returns, that is, an increase in production rate causes an increase in required resources (and hence cost). A negative slope implies increasing returns since an increase in rate requiring an increase in resources decreases unit cost.

As discussed by Cox and Gansler<sup>8</sup>, different signs for  $\beta_2$ , even if statistically significant, are not necessarily contradicting. In the short-run, both increasing and decreasing returns can exist. Even if the data indicate falling unit cost as rate increases, this does not necessarily imply increasing returns to the variable factors. The firm could be producing in the region of diminishing returns on the short-run cost surface, but the dominating learning (cumulative quantity) effect could be causing unit costs to decline.

After considering the above cost impacts, this research demonstrates that estimates obtained from production data using equation (2) are often

inconsistent with optimal contractor behavior. In addition, the assertion that the production on all major Air Force programs since the F-100 have been characterized by increasing returns (lower than optimal production rates) is also examined. On the surface this assertion seems illogical if the contract is written so as to induce cost minimizing behavior and the contractor is interested in making a profit. If the assertion were true, the contractor would certainly have incentives to increase production rate.

#### IV. CONTRACTOR BEHAVIOR

The model presented by Womer<sup>14</sup> is used to demonstrate that negative slopes for rate/cost ( $\beta_2 < 0$ ) result in optimal behavior which is inconsistent with observed and logical contractor behavior. Consider the following defining notation:

$C$  = total discounted program cost,

$q(t)$  = production rate at time  $t$ ,

$Q(t)$  = cumulative production at time, i.e.,  $Q(t) = \int_0^t q(\tau) d\tau$ ,

$x(t)$  = the use rate of a variable composite resource at time  $t$ ,

$\gamma$  = a parameter describing the returns to the variable resource,

$\delta$  = a learning parameter,

$\rho$  = the discount rate,

$A$  = a constant,

$V$  = the total planned units to be produced,

$T$  = the planned time horizon for the program.

The following production function is specified:

$$q(t) = Ax^{1/\gamma}(t)Q^\delta(t). \quad (3)$$

For the moment no assumption is made about the sign of  $\gamma$ , but it is assumed that  $A > 0$  and  $0 < \delta \leq 1$ . Notice that solving equation (3) for  $x(t)$  yields an

improvement curve if production rate is assumed constant, i.e.,

$$\frac{x(t)}{q(t)} = A^{-\gamma} q^{\gamma-1}(t) Q^{-\delta}(t). \quad (4)$$

Also, notice the equivalence with equation (2) when production rate is allowed to vary. Note that  $\beta_1 = -\delta\gamma$  and  $\beta_2 = \gamma - 1$ . The relationship in equation (4) is combined with a behavioral assumption to construct a simple model of optimal firm behavior.

There may be some discussion about the appropriate assumption that governs the firm's behavior. The assumption here is that the contract is structured so as to induce cost minimizing behavior on the part of the contractor. This could be in the form of a fixed price contract, but most likely as a cost-plus incentive or award fee contract.

If cost is measured in units of the variable resource, the firm's objective may be stated as

$$\text{Minimize } C = \int_0^T x(t) e^{-\rho t} dt \quad (5)$$

subject to:

$$q(t) = A x^{1/\gamma}(t) Q^{\delta}(t),$$

$$Q(0) = 0,$$

$$Q(T) = V,$$

$$x(t) \geq 0.$$

This is a problem in optimal control theory, but if it is assumed that the last constraint  $[x(t) \geq 0]$  is satisfied, it is possible to use classical variational techniques to solve the problem. The complete solution to the problem presented in (5) is not needed to demonstrate the hypothesized result, however a transformation simplifies the required analysis somewhat. Let  $Z(t) = Q^{1-\delta}(t)/(1-\delta)$ . This implies  $z(t) = dZ/dt = Q^{-\delta}(t)q(t)$ . The optimization problem may now be restated as

$$\text{Minimize } C = \int_0^T A^{-\gamma} z^\gamma(t) e^{-\rho t} dt \quad (6)$$

subject to:

$$z(0) = 0$$

$$z(T) = v^{1-\delta}/(1-\delta).$$

A sketch of the calculus of variations solution is as follows. Let

$$I = \int_0^T [z_o(t) + \epsilon h(t)]^\gamma e^{-\rho t} dt \quad (7)$$

where  $h(t)$  is a function that gives the difference between the assumed optimal path,  $z_o(t)$ , and any other path, and  $\epsilon$  is an arbitrary constant (see Kamien and Schwartz<sup>11</sup> for a comprehensive derivation]. If equation (7) is treated as function of  $\epsilon$ , the optimum must occur when  $\epsilon=0$ . The derivative is

$$I'(\epsilon) = \int_0^T \gamma [z_o(t) + \epsilon h(t)]^{\gamma-1} h(t) e^{-\rho t} dt, \quad (8)$$

and after equating with zero the following is obtained:

$$I'(0) = 0 = \int_0^T \gamma z_o^{\gamma-1}(t) h(t) e^{-\rho t} dt. \quad (9)$$

After integrating equation (9) by parts, it is possible to obtain the Euler equation of the calculus of variations. Womer<sup>14</sup> has derived the extremals for production rate, resource use rate, and discounted cost by solving the Euler equation.

Additional insight is gained by examining the second variational. This is stated as

$$I''(\epsilon) = \int_0^T \gamma(\gamma-1) [z_o(t) + \epsilon h(t)]^{\gamma-2} h^2(t) e^{-\rho t} dt. \quad (10)$$

After evaluating at zero, the following expression is obtained:

$$I''(0) = \int_0^T \gamma(\gamma-1) z_o^{\gamma-2}(t) h^2(t) e^{-\rho t} dt. \quad (11)$$

For the problem in (6) to have a minimum, it is necessary that (11) be

nonnegative. As long as  $\gamma > 1$ , equation (11) is nonnegative and the solution is a minimum. From an economic point of view, it is obvious what type of behavior is implied when  $\gamma < 1$ . The contractor has incentive to delay all production to the end of the program because of the combined efforts of increasing returns and discounting. It is easy to show mathematically that you can make the integral in (6) approach zero by letting  $z(t)$  be zero until the last instance of time. In short, the solution does not make sense. It implies contractor behavior that is inconsistent with observed contractor behavior.

The author does not deny that increasing returns ( $\gamma < 1$ ) may exist, particularly during the start-up period of production. However, economic theory suggests that the contractor will add resources if the contract is written to induce such cost minimizing behavior. It is highly unlikely that increasing returns to the variable resources exist throughout the production program. This type of irrational contractor behavior has not been noticed in previous research. Certainly the contractor would not plan (as shown by the model) to be in such a situation after start-up.

There are many applications where estimates of the parameters in equation (2) are provided. For example, Bemis<sup>4</sup> provides a table of estimates for many defense items. The estimated values for the quantity slope ( $\beta_1$ ) and the rate slope ( $\beta_2$ ) are transformed to the corresponding  $\gamma$  and  $\delta$  values as presented in equation (4) and are presented in Table 1. Notice that all of the estimated values for  $\gamma$  are less than one. The estimates for tactical missile programs presented by Cox and Gansler<sup>8</sup> are transformed and presented in Table 2. While the Bullpup and Tow estimates seem reasonable, the Sparrow and Sidewinder estimates have  $\gamma < 1$ .

<u>System</u>	<u><math>\gamma</math></u>	<u><math>\delta</math></u>
Aircraft A	.9635	.4692
Aircraft C	.6690	.5580
Aircraft D	.4436	.4084
Aircraft E	.4265	.3377
Aircraft F	.1966	1.0560
Aircraft G	.7031	.3578
Aircraft H	.8703	.4902
Helicopter	.8367	.3047
Jet Engine A	.8797	.4718
Jet Engine B	.8400	.5786
Missile G&C	.2485	.4904
Ordnance Item A	.9561	.1912
Radar Set A	.8286	.1245
Radar Set B	.8734	.0183

Note: Six additional items were included in Bemis' data summary. These were not included because in four cases the rate slope was not provided, and in two cases the quantity slope was not provided.

Table 1. Values of  $\gamma$  and  $\delta$  implied by Bemis' data summary.

<u>System</u>	<u><math>\gamma</math></u>	<u><math>\delta</math></u>
Sparrow (1st source)	.9782	.2467
Sparrow (2nd source)	.8844	.2197
Bullpup	1.0058	.2794
Tow	1.0101	.0129
Sidewinder	.7119	-.0931

Table 2. Estimates of  $\gamma$  and  $\delta$  for missile programs as presented by Cox and Gansler

## V. DISCUSSION

The combined influence of improvement and production rate on cost is still a topic that requires much additional research. Most previous modeling attempts must be interpreted with extreme care because they suffer from severe statistical problems. If there were no data problems (e.g., engineering change orders) and production rate could be measured accurately, the regression equation may be a valid tool for prediction purposes. However, any attempt to make any statement about the estimates from equation (2) is futile.

In terms of production planning, a contractor certainly would not plan to operate at a less than optimal rate if cost minimization is induced by the contract. In many production programs planned production has exceeded actual production. Only the inability of the contractor to deliver on cost and on schedule has resulted in decreased production. In view of the results of this research, it would seem that the above phenomena could imply that contractors are producing at greater than optimal rates given fixed facilities, so that diminishing returns to the variable resources exist throughout the program after some start-up period. Again it must be noted that this does not imply the average cost per unit would have to rise as output rate is increased. The improvement effect could dominate the rate effect and average cost per unit could decline.

This research suggests that regression models as presented in equation (2) are not the answer to the problem. These results are noted and studied since Air Force Systems Command has developed and most likely will use equations such as those analyzed in this paper.

#### VI. AN ALTERNATIVE RATE VARIATIONS MODEL

After discussing this problem with Aeronautical Systems Division (ASD/ACCR) personnel, it became clear that a pure statistical model was not the proper methodology for solving this problem. In particular, the data was extremely deficient. On some programs (e.g., B1-B), the data on lengthy production runs is just not available. On other programs significant engineering changes caused the data to be unreliable. Also, conversations with Col. L. L. Smith shed additional insight on the problem. Smith noted



the following. The models constructed for Air Force Systems Command (these models use the Bohn and Kratz methodology) are used to predict total flyaway cost for a particular program. The modeling methodology was never intended to examine total flyaway cost. The total flyaway cost problem is a much more difficult problem, a cost accounting problem. Smith calls this problem the "budget" problem. The methodology developed by Smith was meant to be applied in the production phase of an airframe program. It was not designed or deemed capable of predicting total flyaway cost.

The above comments were confirmed by Aeronautical Systems Division (ASD/ACCR). In fact, the cost information that is needed as output from the rate variations model is a proforma cost budget for each program. This budget provides annual cost estimates for various categories for 10 years into the future. For example, the cost breakdowns for the F-15 weapon system program are presented in Table 3.

Airframe	Nonrecurring	Gross Weapon System
Propulsion	Flyaway (total)	Gross Weapon System (U)
Electronics	Flyaway (unit)	Advance Buy Credit
Armament	Training	Net Weapon System
Other	PGSE	Advance Buy
ECO	Data	Total Requirements
Recurring Flyaway (total)	SE Conformal Full Tanks	Initial Spares
Recurring Flyaway (unit)	Peculiar Support Total	Total Procurement

Table 3. Cost Categories required for the F-15 Weapon System program for a ten year time horizon.

Many of the items in Table 3 may be estimated without a model by Air Force personnel with a high level of confidence. But the point is, a single equation statistical model will never be able to achieve the level of disaggregation required by Table 3.

The initial strategy of this research effort was to identify the important factors that impact cost through production rate changes and then

to try to integrate these factors into some type of prediction model. Two things became clear immediately: (1) some of the important cost drivers are qualitative and should be included subjectively, and (2) the redistribution of contractor overhead costs was by far the most important cost driver. In fact, Aeronautical System Division personnel stated that if they could determine the redistribution of overhead costs, they could construct a reasonable consistent estimated cost budget.

Therefore, to reduce the scope of the research effort, the decision was made to construct a model that examines overhead redistribution after production rate changes. Also, because of the size of the project and the severe time limitations, it was decided that only the F-16 program should be examined. This decision was partly for convenience since the F-16 System Program Office is located at the research location. At the time the project was initiated, it was understood that there was no chance of completing the project during the ten week research period. This effort is one component of an ongoing research project.

#### VII. OVERHEAD REDISTRIBUTION MODELS

After examining the relevant literature, one methodology appeared to be clearly superior considering the type of data that is available on weapon systems programs. Balut<sup>3</sup> presents a model for redistributing overhead costs. The model was developed for the Office of the Assistant Secretary of Defense (Program Analysis and Evaluation). The model is not described in detail in this report because of space limitations. The interested reader should consult Balut's<sup>3</sup> original paper. The technique presented by Balut is a two step procedure. The first step requires the construction of a program price curve, at the recurring flyaway level, and repricing according to the

given quantity change. The second step is an adjustment for changes in production rate.

The methodology is extremely data intensive since contractor overhead data is required for each program, however the data is available through the Contractor Cost Data Reporting System. Balut stated in personal conversations that the data analysis is an absolute requirement, but it is worth the effort. He has had extremely good luck with the model on a variety of major weapon systems programs. Still, the model was approached with cautious optimism since the F-16 program is quite different from many other programs. The overhead costs are affected by the fact that General Dynamics does not own the building or the land where the aircrafts are produced. This leads to a situation where the short-run fixed overhead percentage is much lower than the aerospace industry average. In short, some modifications on Balut's methodology were anticipated.

#### VIII. THE DATA

The data was taken from two sources:

1. The overhead data was taken from the annually produced plant wide data summaries (DD Form 1921-3).
2. The quantity data was provided by the F-16 System Program Office.

The compilation of the data into the proper form was a major part of the research effort. In fact, half of the research period was spent collecting and adjusting data. In the last week of the research period, a complete and consistent data base for the F-16 program was finally obtained. The data is not reproduced here since it was submitted in confidence to the government. It is for official use only, and it cannot be disclosed without prior written permission from General Dynamics, Fort Worth Division.

#### IX. PRELIMINARY FINDINGS

The initial results from the modeling effort indicate that the F-16 model will be able to predict overhead redistributions as rate changes. It must be stressed that these results are tentative. Time limitations and software limitations at the research location prevented a complete analysis of the model. The final model that will be presented to the Business Research Management Center will be completed early in the Fall 1984.

#### X. RECOMMENDATIONS

The recommendations are somewhat dependent on the final outcome of the empirical work, however it appears that the following can be stated:

1. The production rate variations problem at the total flyaway level is a cost accounting problem, not a simple statistical problem. It should be recognized as such.
2. The problem is solvable if it is possible to predict overhead cost redistribution after production rate changes. Therefore Air Force Systems Command should supplement its current statistical formulations with models that provide the information that is needed by Aeronautical Systems Division.

It should be noted that the construction of appropriate models will be a major effort. However, a considerable amount of work has been completed. For example, the data is available at the Office of the Assistant Secretary of Defense. The major effort will involve analyzing and organizing the data in the proper form for model construction, but irregardless of the research hours required, the project should be initiated.

## REFERENCES

1. A. A. Alchian. 1959. Costs and outputs. In the Allocation of Economic Resources (M. Abramovitz, et. al., eds.) Stanford: Stanford University Press, 23-40.
2. A. A. Alchian. 1963. Reliability of progress curves in airframe production. *Econometrica*, 31 (October), 679-693.
3. S. J. Balut. 1981. Redistributing fixed overhead costs. *Concepts*, 4, 63-76.
4. J. C. Bemis. 1983. A model for examining the cost implications of production rate. Paper presented at 50th meeting of the Military Operations Research Society, Annapolis, Maryland.
5. J. C. Bemis. 1981. A model for examining the cost implications of production rate. *Concepts*, 4 (Spring), 84-94.
6. M. Bohn and L. A. Kratz. 1984. The impact of production rate on unit costs. (Paper presented at the ORSA/TIMS joint national meeting) San Francisco, California.
7. L. Cox, B. Dembroski, D. Elam, and P. Martin. 1981. Analyses to support evaluation of the potential benefits of competition in cruise missile production. Arlington, Virginia: The Analytic Sciences Corporation.
8. L. W. Cox and J. S. Gansler. 1981. Evaluating the impact of quantity, rate, and competition. *Concepts*, 4 (Autumn), 29-53.
9. R. Crouch, 1980. Avoiding bias in progress functions. *Defense Management Journal*, 16 (Third Quarter), 40-45.
10. B. Gold. 1981. Changing perspectives on size, scale, and returns. *Journal of Economic Literature*, 19 (March), 5-33.
11. M. I. Kamien and N. L. Schwartz. 1981. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. New York: North-Holland.
12. J. P. Large, K. Hoffman, and F. Kontrovich. 1976. *Production Rate and Production Costs*, R-1609-PA&E Santa Monica: The Rand Corporation.
13. L. L. Smith. 1976. *An Investigation of Changes in Direct Labor Requirements Resulting from Changes in Airframe Production Rate*, PhD. Thesis. Eugene: The University of Oregon.
14. N. K. Womer. 1979. Learning curves, production rate, and program costs. *Management Science*, 25 (April), 312-319.

15. N. K. Womer and T. R. Gullledge, Jr. 1983. A dynamic cost function for an airframe production program. *Engineering Costs and Production Economics* (September), 213-27.
16. T. P. Wright. 1936. Factors affecting the cost of airplanes. *Journal of Aeronautical Sciences*, 3 (February), 122-128.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THERMAL STABILITY CHARACTERISTICS OF SILAHYDROCARBONS

Prepared by: Dr. Vijay K. Gupta, and Dennis W. Weatherby

Academic Rank: Associate Professor

Department and University : Chemistry Department  
Central State University  
Wilberforce, Ohio 45384

Research Location: Materials Laboratory (AFWAL/MLBT)  
Wright Patterson Air Force Base, Ohio

USAF Research: Mr. C.E. Snyder, Jr.

Date: August 31, 1984

Contract No: F 49620-82-C-0035

# THERMAL STABILITY CHARACTERISTICS OF SILAHYDROCARBONS

by

Vijay K. Gupta and Dennis W. Weatherby

## Abstract

The silahydrocarbon class of materials were selected for investigation to provide candidate fluids usable over the temperature range of  $-54^{\circ}\text{C}$  to  $315^{\circ}\text{C}$ . Thermal and hydrolytic stability of these fluids were the two main areas of concern. The following four silahydrocarbons chosen for investigation were:  $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3$ ,  $\text{CH}_3\text{Si}(\text{C}_8\text{H}_{17})_2(\text{CH}_2)_3\text{CH}(\text{CH}_3)_2$ ,  $\text{C}_2\text{H}_5\text{Si}(\text{C}_8\text{H}_{17})_3$  +  $\text{C}_2\text{H}_5\text{Si}(\text{C}_{10}\text{H}_{21})_3$ , and  $(\text{CH}_3)_2\text{SiCH}(\text{CH}_3)_2\text{C}_{14}\text{H}_{29}$ . The thermal stability studies of the above silahydrocarbons were conducted with respect to temperature, the heating time, the moisture content, and the nature of the metal container. It appears that silahydrocarbons  $(\text{CH}_3)_2\text{SiCH}(\text{CH}_3)_2\text{C}_{14}\text{H}_{29}$  and  $\text{CH}_3\text{Si}(\text{C}_8\text{H}_{17})_2(\text{CH}_2)_3\text{CH}(\text{CH}_3)_2$  are slightly more stable as compared to the other two silahydrocarbons. The branching of one of the substituent group seems to have contributed to the added thermal stability. More than 90 percent of the silahydrocarbon  $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3$  was decomposed when heated at  $398.9^{\circ}\text{C}$  for 6 hours, and there was little or no decomposition of the above hydrocarbon until the temperature of  $329.8^{\circ}\text{C}$ . The presence of 0.2% moisture and the metal container of the bomb material also have no significant impact on the thermal decomposition characteristics.



#### Acknowledgements

The authors would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing them with the opportunity to spend a worthwhile and interesting summer at the Materials Laboratory, Wright Patterson Air Force Base, Dayton, Ohio 45433. We would also like to acknowledge the laboratory, in particular the Non-structural Materials Branch, for its hospitality and excellent working conditions.

Finally, we would like to thank Mr. Carl E. Snyder Jr., for suggesting this area of research and for his collaboration and helpful suggestions. We would also like to acknowledge many helpful discussions with Ms. Lois Gschwender, and the technical assistance provided by the University of Dayton Research Institute's Lubricants Branch. The authors are thankful to Mr. Jim Munta of the Analytical Services Branch and Dr. Jeff Workman, and Dr. Chin Yu of Technical Services Inc. for GC/MS analysis.

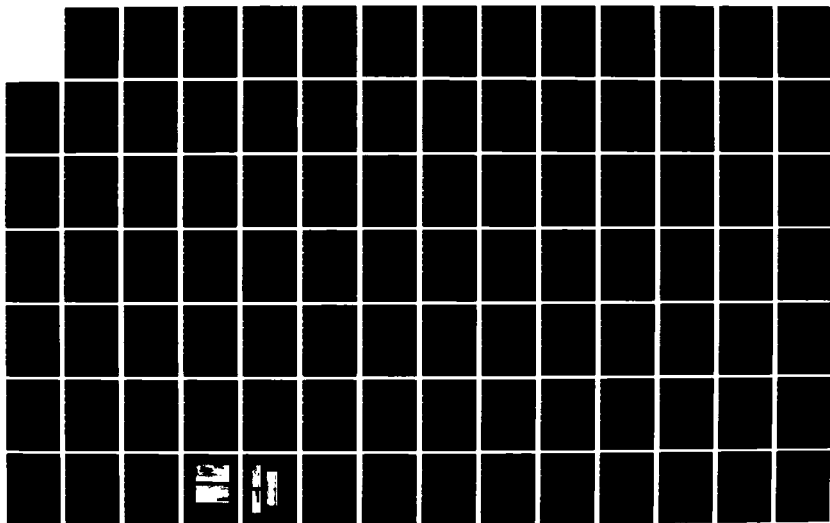
## I. INTRODUCTION

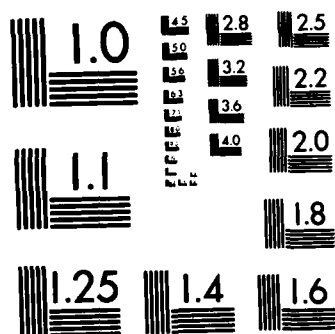
Recent effort has led to the development of many functional fluids for use in aerospace applications such as jet engine oils, greases, and hydraulic fluids. Since hydraulic fluids are not expected to operate in oxidative environments, therefore thermal and hydrolytic stability are the two main areas of concern (1). In order to meet the need for a high temperature fluid, MIL-H-27601, a - 40°C to 288°C hydraulic fluid was developed. A highly refined, deep-dewaxed paraffinic mineral oil was selected as the base stock for the above application, and the fluid provided satisfactory service for over ten years in the temperature range of - 40°C to 288°C.

The development of supersonic missiles, like the advanced strategic air launched missile (ASALM) has created the need for a hydraulic fluid useable over the temperature range of - 54°C to 315°C. The mineral oil base stocks from which MIL-H-27601 is derived, do not have the needed viscosity temperature properties to perform satisfactorily in the temperature range of - 54°C to 315°C. Synthetic hydrocarbons based on hydrogenated poly-alphaolefin oligomers have also been found to be deficient both in viscosity temperature properties and thermal stability (2). Perfluorinated fluids which have excellent thermal and oxidative stabilities, have several disadvantages such as high density, poor bulk modulus, elastomer incompatibility, and lack of suitable additives (3).

In response to the need for hydrocarbon-type fluids with improved properties, a Materials Laboratory Program has led to the development of a class of compounds called silahydrocarbons. These compounds have excellent viscosity temperature properties and thermal stability and are expected to

AD-A154 337 UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 3/13  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.  
UNCLASSIFIED W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480 F/G 5/1 NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

be hydrocarbon like in their physical and chemical properties. Therefore, silahydrocarbons are good candidates for hydraulic fluids useable over the - 54°C to 315°C temperature range.

The silahydrocarbons are defined as compounds of silicon substituted by four hydrocarbon groups e.g.,  $\text{SiR}_4$ . In general, R can be any hydrocarbon group such as  $\text{CH}_3$ ,  $\text{C}_2\text{H}_5$  -  $\text{C}_{12}\text{H}_{25}$  and be either primary, secondary, tertiary, or a branched alkyl group. The silahydrocarbon may be synthesized as symmetrical  $\text{SiR}_4$ , e.g.,  $\text{Si}(\text{CH}_3)_4$  where all four R groups are identical or unsymmetrical  $\text{RSiR}'_3$ , e.g.,  $\text{CH}_3\text{Si}(\text{C}_{12}\text{H}_{25})_3$  where R and R' groups are different. In the unsymmetrical structures variations can be made at will by proper selection of the R groups. The properties such as vapor pressure, boiling point, viscosity, and thermal stability can be controlled by the proper selection of the substituent groups(R). For example the length of the R group effects the boiling point,  $(\text{CH}_3)_4\text{Si}$  has a boiling point of 26°C while  $(n\text{-C}_{12}\text{H}_{25})_4\text{Si}$  has a boiling point of 250°C at 0.06 mm. One expects some dependence of thermal stability on the structure and size of the substituent groups in the silahydrocarbon. This report describes the effort to understand the thermal stability characteristics of silahydrocarbons as a function of the variation in the structure of the substituent group.

## II. OBJECTIVES

The main objective of the summer program was to understand the thermal stability characteristics of silahydrocarbons as a function of temperature, and time in the dry environment. The other objectives were to study the effect of moisture, the effect of metal catalyst, and to understand the

impact of branching of the substituent group in the silahydrocarbon on the thermal stability characteristics.

### III, PROPOSED EFFORT

It was proposed to study the thermal stability characteristics of the following four hydraulic fluids which were assigned the MLO numbers as indicated below:

1. MLO 84-193  $\text{CH}_3 \text{Si} (\text{C}_{10}\text{H}_{21})_3$
2. MLO 84-276  $\text{CH}_3 \text{Si} (\text{C}_8\text{H}_{17})_2 (\text{CH}_2)_3 \text{CH} (\text{CH}_3)_2$
3. MLO 84-284  $\text{C}_2\text{H}_5 \text{Si} (\text{C}_8\text{H}_{17})_3 + \text{C}_2\text{H}_5 \text{Si} (\text{C}_{10}\text{H}_{21})_3$
4. MLO 84-285  $(\text{CH}_3)_2 \text{CH Si}(\text{CH}_3)_2 \text{C}_{14}\text{H}_{29}$

The oils MLO 84-193 and MLO 84-284 have straight chain substituent groups, whereas the oils MLO 84-276 and MLO 84-285 have one branched alkyl group. The oils MLO 84-276, MLO 84-284, and MLO 84-285, provided by Gulf Research and Development Company, were used as received without any purification, since the quantities of these oils were small. The oil MLO 84-193 was purified, and the purification is described in the experimental section. Thermal studies of the above oils were proposed with respect to temperature, the heating time, the moisture content, and the nature of the metal container. The above studies were intended to explore the impact of branching of the substituent group in the silahydrocarbon on the thermal stability characteristics.

### IV, EXPERIMENTAL

This section describes the purification of the oil MLO 84-193, and the test procedures.

Purification: The major component of the oil MLO 84-193 is  $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3$ .

and the oil was purified using Spinning-band distillation column under vacuum. The earlier fractions did remove the low boiling components and most of the impurities, and the middle fraction in the temperature range of 190-195°C at 0.5 mm Hg pressure was collected and labeled as MLO 84-193 Distillate 6. The fraction distillate 6 was used for all the thermal stability studies. The purification of distillate 6 was checked using GC and GC/MS analysis. The chromatogram of an earlier fraction called distillate 1 is given in Figure 1, and the data in table 1 lists the various impurities found in MLO 84-193. Figure 2 is the chromatogram of distillate 6 of the above oil indicating that the distilled fraction contained the following components as listed below:

$\text{CH}_3\text{SiC}_8\text{H}_{17}(\text{C}_{10}\text{H}_{21})_2 = 0.3\%$   
 $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3 = 96.4\%$   
 $\text{CH}_3(\text{C}_{10}\text{H}_{21})_2\text{SiO}(\text{C}_{10}\text{H}_{21})_2\text{CH}_3 = 2.6\%$   
 Other impurities = 0.7%

whereas in the oil before distillation the component  $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3$  was about 79.6% and remaining impurities.

Thermal Stability Test Procedure: The bombs were made of 8" long stainless steel tubes either 0.25" or 0.375" diameter with swagelok fittings on both ends. The bombs were cleaned with a suitable solvent such as hexane and dried in an oven at 100°C for one hour. Approximately 2.0 cc of the test fluid was added to the bomb, and the lamp grade nitrogen was bubbled through the fluid to remove any air and the bomb was quickly capped with swagelok fitting. The bomb was then placed in an oven at the specified temperature (controlled within 2°C for a specific time period. After the bomb was heated in the oven for a given length of time, the bomb was removed, allowed to cool, and was placed in a bath maintained at -54°C for at least 30 minutes in order for the

gaseous components to liquify or freeze. The bomb was taken out of the cold bath, and one of the cap was replaced by a cap with a small hole containing a septum in order to facilitate the removal of sample for gas analysis by GC. The method and gas chromatographic conditions for the gaseous analysis are listed in Table 2. After the gas phase sample called HEADSPACE was analyzed the bomb was opened and the liquid sample was transferred to a glass vial and analyzed for liquid components. The method and gas chromatographic conditions are listed in Table 3. All the four hydraulic fluids were studied for thermal stability characteristics using the above test procedure, and the results are presented in the next section.

#### V. RESULTS AND DISCUSSION

All the four hydraulic fluids MLO 84-193 Distillate 6, MLO 84-276, MLO 84-284, and MLO 84-285 were studied as a function of time at 371.1°C ( 700 F ). The data in table 4 gives the concentrations of the four oils as function of time 0 to 24 hours. These concentrations represent the amount of specific component determined from GC Analysis when the oil is stressed at 371.1°C for a specific time. The zero hour means that the oil is not stressed at all. From the data in the table it appears that the oil MLO 84-285 is definitely more stable as compared to other oils, but in case of the oil MLO 84-276, this was significantly impure. If a correction is made for the amount of impurity, MLO 84-276 may exhibit better stability. The data in table 5 gives the viscosities of the four oils as a function of temperature. The oils MLO 84-193 and MLO 84-285 become solid at - 40°C or below, whereas the oils MLO 84-276, and MLO 84-284 have suitable viscosities in the temperature range of -54°C to 98.9°C. The changes in viscosities of the oils MLO 84-193, MLO 84-276, MLO 84-284, and MLO 84-285 at 37.8°C, when



stressed at 371.1°C for 6 hours are - 16.6%, -8.3%, -14.3% and -11.8% respectively. From the above viscosity changes, it appears that the oils MLO 84-276 and MLO 84-285 are relatively more thermally stable as compared to oils MLO 84-193 and MLO 84-284. Both from the thermal stability and viscosity data in tables 4 and 5 respectively, it appears that the branching of one of the substituent group in the oils MLO 84-276 and MLO 84-285 may have contributed to the stability of the oils.

The effect of temperature on the decomposition of MLO 84-193 distillate 6 was studied and the stress time in each case was fixed as six hours, and the data is given in table 6. From the table it appears that there is hardly any decomposition up to a temperature of 329.8°C as indicated by the GC analysis of the liquid. From 329.8°C to 398.9°C, the rate of decomposition increased so fast that at 398.9°C only about 3.8% of the  $\text{CH}_3\text{Si}(\text{C}_{10}\text{H}_{21})_3$  component was left undecomposed. The GC analysis of the gaseous components in the headspace was the only indication of decomposition of the oil in the temperature range of 287.8°C to 329.8°C.

The data in tables 7 to 14 and figures 3 to 10 represent the typical examples of the decomposition fragments as identified by GC/MS analysis, and the chromatograms of the four oils in gaseous and liquid phases. It is a general observation that the gaseous phase contains various hydrocarbons from one carbon atom to 10 carbon atoms. With the exception of methane, the hydrocarbons identified are either saturated hydrocarbon or a hydrocarbon containing one double bond or a cyclic hydrocarbon. In addition some low molecular weight silahydrocarbons have also been identified. In order to seek better understanding of the decomposition pattern of the various substituents groups in the silahydrocarbons, the thermal decomposition of n-octane, n-decane, n-dodecane, and n-tetradecane was investigated at

371.1°C for 6 hours stress time. The headspace analysis indicated the gaseous fragments similar to the gaseous fragments as found in the case of silahydrocarbons with the exception that no low molecular weight silahydrocarbon peaks were identified. In the liquid phase the hydrocarbons with 5 or more carbon atoms have been identified along with various silahydrocarbons. It is generally observed that the hydrocarbon with no. of carbon atoms same as the no. of carbon atoms in the substituent group of the silahydrocarbon was found in greater concentration than other hydrocarbons. For example in the case of oil MLO 84-193 distillate 6 which contained  $C_{10}H_{21}$  groups, the hydrocarbons with ten carbon atoms were present in greater concentration as compared to other hydrocarbons. Similar observation was made in the case of other oils such as in MLO 84-285 which contains  $C_{14}H_{29}$  group contained more of hydrocarbons with 14 carbon atoms, and MLO 84-276 which contained  $C_8H_{17}$  group more of Hydrocarbons with 8 carbon atoms were found. Based on the above observation, it may be assumed that the alkyl group in the silahydrocarbon is broken at Si - C bond, if that is so one would expect higher concentration of the corresponding silahydrocarbon. For example, let us take the case of  $CH_3 Si (C_{10}H_{21})_3$ , when one of the  $C_{10}H_{21}$  group comes off, it will leave the structure  $H Si CH_3(C_{10}H_{21})_2$ . Thus when  $C_{10}H_{20}$  is found in abundance, one will also expect  $H Si CH_3(C_{10}H_{21})_2$  to be in abundance too. On examining the data in table 11, it is observed that fragments  $H Si CH_3(C_{10}H_{21})_2$  and  $(CH_3)_2 Si (C_{10}H_{21})_2$  are more abundant than other decomposition fragments. Since the hydrocarbon with 9 carbon atoms is not formed in abundance, the bond between first and second carbon atoms of the alkyl group is relatively more stable than the bond between Si and Carbon atom. One may argue that the silahydrocarbon  $H Si CH_3 (C_{10}H_{21})_2$

is less stable than the hydrocarbon.  $(CH_3)_2 Si (C_{10}H_{21})_2$ , and picks up a carbon atom to form the second compound. However, the silahydrocarbons containing Si-H bond appear to be stable enough to be detected in an infra red spectra of the post-stress sample. A band corresponding to Si-H bond was observed in all the four hydraulic fluids at  $2050\text{ cm}^{-1}$ , after they were stressed at  $700F(371.1^\circ C)$  for 8 hours, a similar band was not found in the infra red spectra of the unstressed hydraulic fluids. Based on the above arguments, it appears that the Si-C bond may be slightly weaker than other C-C bonds of the alkyl groups specially when the silahydrocarbon contains a long chain alkyl group as a substituent group. In case of MLO 84-276 which contains groups with 8 or less carbon atoms, the band for Si-H bond was a very weak band indicating that Si-C bond is not easily broken. In general, it appears that the alkyl group in the silahydrocarbon is broken randomly through the chain length of the alkyl group resulting in all possible silahydrocarbon fragments.

In addition, the following experiments were performed with MLO 84-193 distillate 6 to explore the effect of conditions given below when the fluid is stressed at  $371.1^\circ C$  for 6 hours:

<u>Conditions</u>	<u>% Concentration</u>
The fluid was placed in the stainless steel bomb.	77.0
The fluid was placed in the pyrex glass tube, and the tube was placed in the bomb	81.9
0.2% moisture was added to the bomb along with the fluid.	77.7
Nitrogen gas was not bubbled through the fluid.	80.2

The concentration values listed above represent the concentration of

the  $\text{CH}_3 \text{ Si } (\text{C}_{10}\text{H}_{21})_3$  component that did not decompose during stress. From the results on the previous page, it may be concluded that the presence of 0.2% moisture, the stainless steel container, and not bubbling nitrogen through the fluid did not have any significant impact on the thermal stability of the hydraulic fluid.

#### VI. CONCLUSIONS

On the basis of the results presented in the previous section, the conclusions are as listed below:

- The oils MLO 84-285 and MLO 84-276 appear to be slightly more thermally stable than oils MLO 84-193 and MLO 84-284.
- Both temperature and time have significant impact on the thermal stability. Higher the temperature more the stress, and longer the time of stress more thermal decomposition. However these fluids do require some minimum energy before they start decomposing.
- The branching of the substituent group in the silahydrocarbon seems to increase the thermal stability specially when branching in the substituent group takes place closer to the Si atom as in the case of MLO 84-285.
- The presence of 0.2% moisture in test fluid did not effect the thermal stability.
- The stainless steel from which the bombs are made does not seem to have any catalytic effect on the thermal stability.

## VII. RECOMMENDATIONS

On the basis of our effort and the results accomplished, it is felt that future effort should emphasise the following:

1. The thermal stability bomb tests should be performed quantitatively, i.e., the bombs should be weighed before and after stress.
2. The use of glass liners during the stress conditions may be desirable, since the metal of the bomb material does not seem to catalyze the thermal decomposition reaction.
3. The use of an internal standard during the GC analysis is recommended in order to achieve absolute concentrations of the components rather than the relative concentrations.
4. In order to reach the goal of thermally stable silahydrocarbons, it is recommended that silahydrocarbons with two long chain alkyl groups and two short chain branched alkyl groups should be synthesized and characterized. A typical example can be  $((CH_3)_2CH)_2 Si(C_{10}H_{21})_2$ . The arrangement as suggested above will provide two long chain groups and two short chain branched groups.

RIC  
 07/11/84 13:43:00  
 SAMPLE: MLD84-193#1  
 COMES.: GC-EI  
 RHINE: G 1.1003 LABEL: N 0. 4.0 QUANT: A 0. 1.0 J 0 BASE: U 20, 3  
 180  
 100.0  
 49 109 200 6:40  
 330 442 498 528 611 676 735 774 851 935 1000 SCAN  
 33:20 TIME  
 714752.

Fig. 1: Chromatogram of MLO 84-193 Fraction 1.

FID  
 07/13/84 14:06:00  
 SAMPLE: M084-193-DIT 6  
 COND.: GC-EI  
 RANGE: C 1.1012 LABEL: N 0, 4.0 QUAN: A 0, 1.0 J 0 BASE: U 20, 3  
 112640.

DATA: 2734E #492  
 CALL: CAL13 #3  
 SCANS 1 TO 2000

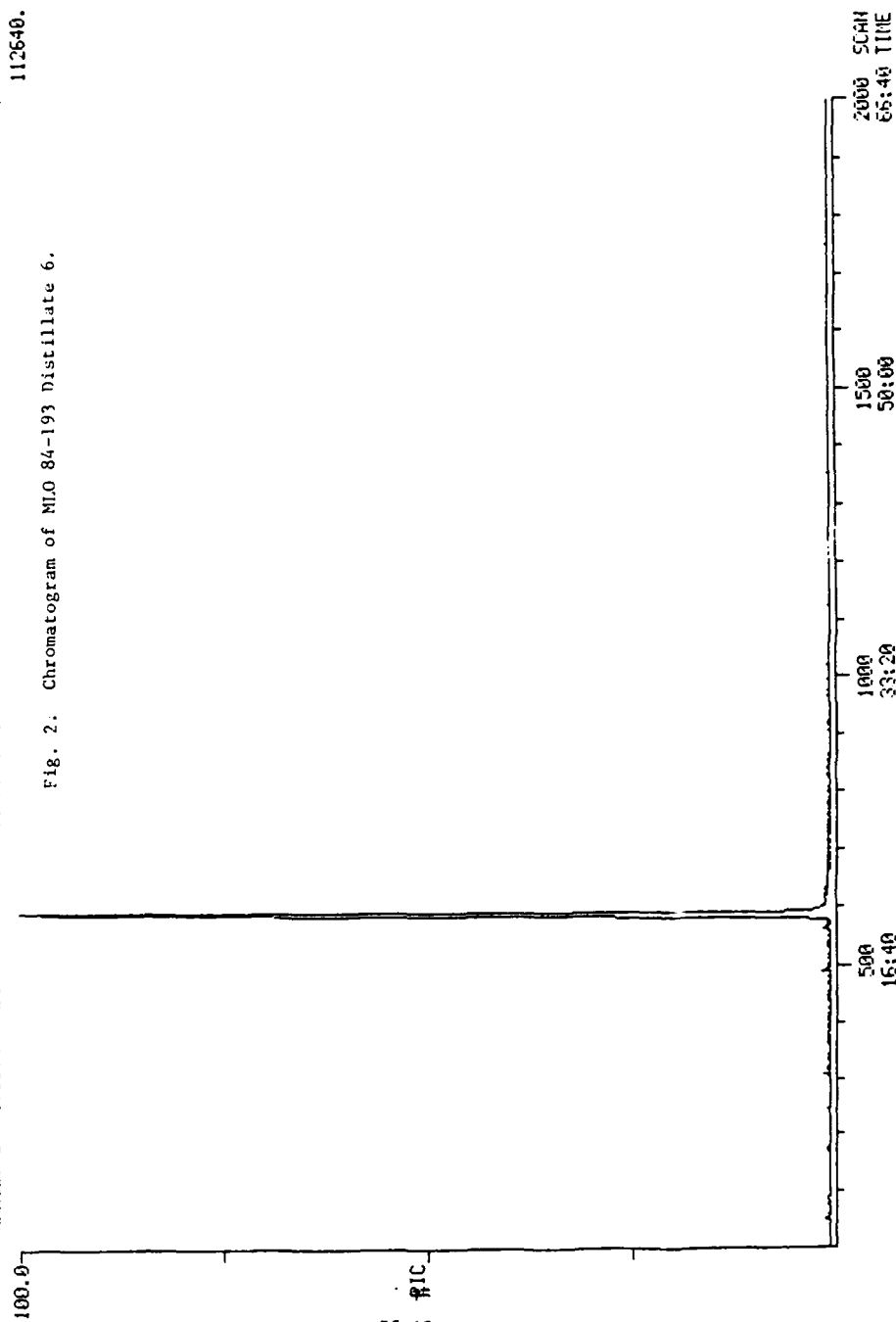


Fig. 2. Chromatogram of MLO 84-193 Distillate 6.

Table 1 GC/MS Analysis of MLO 84-193 Distillate 1.

<u>Scan #</u>	<u>Component</u>
162	H <sub>2</sub> Si(C <sub>3</sub> H <sub>7</sub> ) <sub>2</sub> *
180	C <sub>10</sub> H <sub>21</sub> OH or C <sub>10</sub> H <sub>20</sub>
198	C <sub>10</sub> H <sub>21</sub> OH or C <sub>10</sub> H <sub>20</sub>
207	C <sub>10</sub> H <sub>21</sub> Br
321	C <sub>14</sub> H <sub>27</sub> OH or C <sub>14</sub> H <sub>26</sub>
330	C <sub>14</sub> H <sub>29</sub> OH or C <sub>14</sub> H <sub>28</sub>
350	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>4</sub> H <sub>9</sub> C <sub>10</sub> H <sub>21</sub> * or its isomer
361	C <sub>10</sub> H <sub>21</sub> Br
442	C <sub>20</sub> H <sub>42</sub>
490	HSi CH <sub>3</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>
498	(CH <sub>3</sub> ) <sub>2</sub> Si(C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>
676	CH <sub>3</sub> Si C <sub>8</sub> H <sub>17</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>
774	CH <sub>3</sub> Si (C <sub>10</sub> H <sub>21</sub> ) <sub>3</sub>

\* Refers to the possible components or their isomer.



Table 2 Method and Gas Chromatographic Conditions for the analysis of Gases.

METHOD: CAR6  
CHANNEL 4

1. DATA INPUT

RUNTM #PKS  
40.00, 100

MV/MIN DELAY MIN-AR BUNCH  
.300, 0.00, 1000, NO

INTEGRATOR EVENTS  
TIME EVENT  
1 /E

CONTROL EVENTS  
TIME EVENT ECM RLY  
1 /E

2. DATA ANALYSIS

PROC RPRT SUP-UNK  
ZERO, ME, NO

UNITS TITLE  
AREA %

3. USER PROGRAMS

POST-ANAL DIALG-PRG PARAM-FILE  
/N

4. REPORTS

RDVC #RPTS  
1 T2, 1  
2 /E

DONE

GAS CHROMATOGRAPHIC CONDITIONS

Model: Perkin.- Elmer 900  
Fused Silica Capillary Column  
Length: 35 M  
Diameter: 0.22 MM  
Liquid Phase: Methyl Silicone Carbowax Deactivated  
Split Ratio: 100 to 1  
Aux. Gas: 40 ml/min Carrier Gas: 1 ml/min  
Chart Speed: 1 cm/min  
Detector: FID  
Attenuation: 10 x 16  
Temperatures: Injector: 220°C; Detector: 250°C  
Column Temperature: -50°C to 200°C Program Rate: 8 deg/min  
Initial Hold: 0 min Final Hold: 20 min  
Sample Size: 2 ml  
Sample ID: Heatspace  
Date: 8/24/84

Table 3: Method and Gas Chromatographic conditions for the analysis of liquids.

METHOD: DVK  
CHANNEL 2

1. DATA INPUT

RUNTM #PKS  
60.00, 200

MV/MIN DELAY MIN-AR BUNCH  
.300, 0.00, 10, NO

INTEGRATOR EVENTS  
TIME EVENT  
1 /E

CONTROL EVENTS  
TIME EVENT ECM RLY  
1 80.00, CY  
2 /E

2. DATA ANALYSIS

PROC RPRT SUP-UNK  
ZERO, ME, NO

UNITS TITLE  
AREA %

3. USER PROGRAMS

POST-ANAL DIALG-PRG PARAM-FILE  
HPGC

4. REPORTS

RDVC #RPTS  
1 T2, 1  
2 /E

GAS CHROMATOGRAPHIC CONDITIONS

MODEL:HP5710A  
FUSED SILICA CAPILLARY COLUMN  
LENGTH:12 M  
DIAMETER:0.22 MM  
LIQUID PHASE:METHYL SILICONE CARBOWAX DEACTIVATED  
SPLIT RATIO:100 TO 1  
AUX. GAS:40 ML/MIN CARRIER GAS:1 ML/MIN  
CHART SPEED:20CM/HR  
DETECTOR:FID  
ATTENUATION:10 X 16  
TEMPERATURES:  
INJECTOR:300 DEG.C  
DETECTOR:350 DEG.C  
COLUMN TEMP:70-270 DEG.C PROGRAM RATE:8 DEG/MIN  
INITIAL HOLD:2 MIN FINAL HOLD:32 MIN  
SAMPLE SIZE:0.01  
SAMPLE ID:HYDFL  
DATE:8/24/84  
ANALYST:DVK

Table 4 : Concentrations of the four Hydraulic Fluids as a Function of time at 700°F(371.1°C).

Time(hours)	Concentration of the Silahydrocarbon			
	MLO 84-193 distillate &	MLO 84-276	MLO 84-284	MLO 84-285
0	96.40	82.98	97.68	97.27
4	86.94	-	-	-
5	82.23	-	-	-
6	77.00	67.75	76.40	89.24
7	63.74	61.49	70.14	88.54
8	63.91	55.80	69.63	78.71
16	39.70	42.17	31.86	65.05
24	23.76	20.82	33.95	62.27

Table 5: Viscosities of the four Hydraulic Fluids as a function of Temperature

Hydraulic Fluid	Viscosities <sup>a</sup> at Different Temperatures °C				
	98.9	37.8	- 40	-54	-54 (soak) 37.8 <sup>b</sup>
MLO 84-193	3.53	13.60	solid	solid	10.34
MLO 84-276	2.09	6.85	419.25	2285	1938* 6.28
MLO 84 284	3.03	11.29	664.09	2584	Frozen** 9.68
MLO 84-285	1.94	5.67	solid	solid	5.00

a Viscosities are expressed in centi Stokes.

b Viscosity data after heating the oil for 6 hours at 371.1°C.

\* Each successive run viscosity decreased with time, 1790 after 48 hours soak period.

\*\* The sample was frozen after 24 hours.

Table 6 Effect of Temperature on the Decomposition of MLO 94-193  
Distillate 6. for 6 Hours Stress Time.

Temperature °F (°C)	Concentration of the Silahydrocarbon ( % )
Unstressed	96.4
550(297.8)	96.0
580(304.5)	94.7
600(315.6)	94.7
625(329.8)	94.7
650(343.4)	89.8
675(357.3)	90.7
700(371.1)	77.0
725(385.0)	52.6
735(390.0)	23.5
750(399.9)	3.8

Table 7 : GC/MS Analysis of Gaseous Components of MLO 84-193 Distillate 6  
Stressed at 371.1°C for 7 Hours.

<u>Scan No.</u>	<u>Component</u>	<u>% Concentration</u>
16	CH <sub>4</sub>	32.26
20	C <sub>2</sub> H <sub>6</sub>	29.78
40,42	C <sub>3</sub> H <sub>6</sub> and C <sub>3</sub> H <sub>8</sub>	17.77
50,52	C <sub>4</sub> H <sub>6</sub> and C <sub>4</sub> H <sub>10</sub>	1.03
70	(CH <sub>3</sub> ) <sub>4</sub> Si	0.20
85,93 99,106	C <sub>5</sub> H <sub>10</sub> and C <sub>5</sub> H <sub>12</sub>	8.93
170	(CH <sub>3</sub> ) <sub>3</sub> Si C <sub>2</sub> H <sub>5</sub>	
187,200,206,215,222	C <sub>6</sub> H <sub>12</sub> and C <sub>6</sub> H <sub>14</sub>	3.87
292,304,317	C <sub>7</sub> H <sub>14</sub> and C <sub>7</sub> H <sub>16</sub>	2.88
383,393 396	C <sub>8</sub> H <sub>16</sub> and C <sub>8</sub> H <sub>18</sub>	0.78
463,472	C <sub>9</sub> H <sub>18</sub> and C <sub>9</sub> H <sub>20</sub>	0.12
535 545 550	C <sub>10</sub> H <sub>20</sub> and C <sub>10</sub> H <sub>22</sub>	0.30

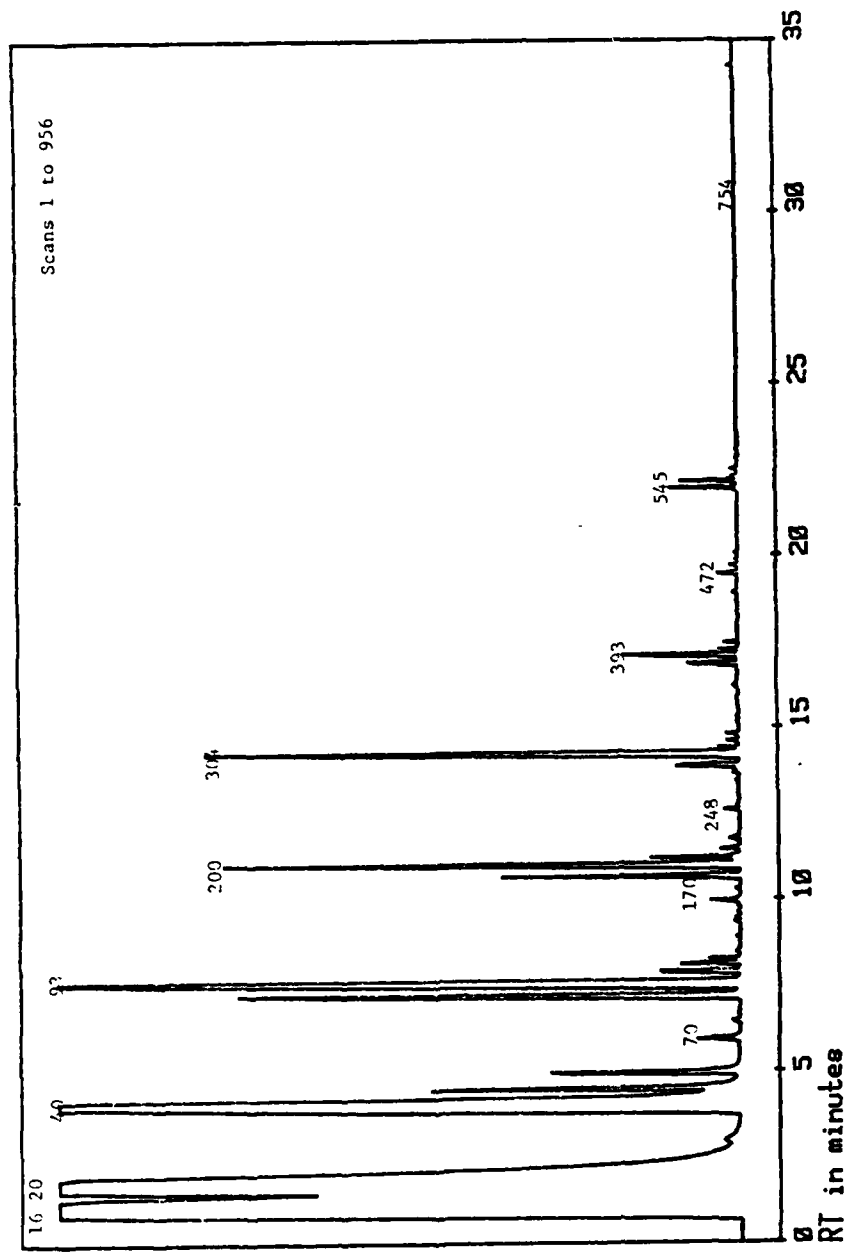


Fig. 3 CHROMATOGRAM MLO 84-193 DISTILLATE, 700F 7 HOURS HEADSPACE

Table 8 GC/MS Analysis of Gaseous Components of MLO 84-276 Stressed at 371.1°C for 6 hours.

Scan No	Component	% Concentration
35	CH <sub>4</sub>	22.49
36	C <sub>2</sub> H <sub>6</sub>	28.01
38	C <sub>3</sub> H <sub>6</sub>	3.07
44	C <sub>4</sub> H <sub>10</sub>	14.44
47	C <sub>4</sub> H <sub>8</sub>	5.06
65	C <sub>5</sub> H <sub>10</sub>	2.28
82, 94	C <sub>5</sub> H <sub>10</sub>	1.12, 3.15
90, 92	C <sub>5</sub> H <sub>12</sub>	2.46, .25
146	C <sub>6</sub> H <sub>12</sub>	3.55
161	C <sub>6</sub> H <sub>12</sub> and C <sub>6</sub> H <sub>14</sub>	1.38
165	C <sub>6</sub> H <sub>12</sub>	0.45
168	(CH <sub>3</sub> ) <sub>3</sub> Si C <sub>2</sub> H <sub>5</sub>	0.81
186	C <sub>6</sub> H <sub>12</sub>	0.73
199	C <sub>6</sub> H <sub>14</sub>	0.33
205, 208	C <sub>6</sub> H <sub>12</sub>	0.30, 0.03
292	C <sub>7</sub> H <sub>14</sub>	0.47
303	C <sub>7</sub> H <sub>16</sub>	0.20
384	C <sub>8</sub> H <sub>16</sub>	0.06
779	C <sub>9</sub> H <sub>20</sub>	0.06



AMPLITUDE X 2

56-25

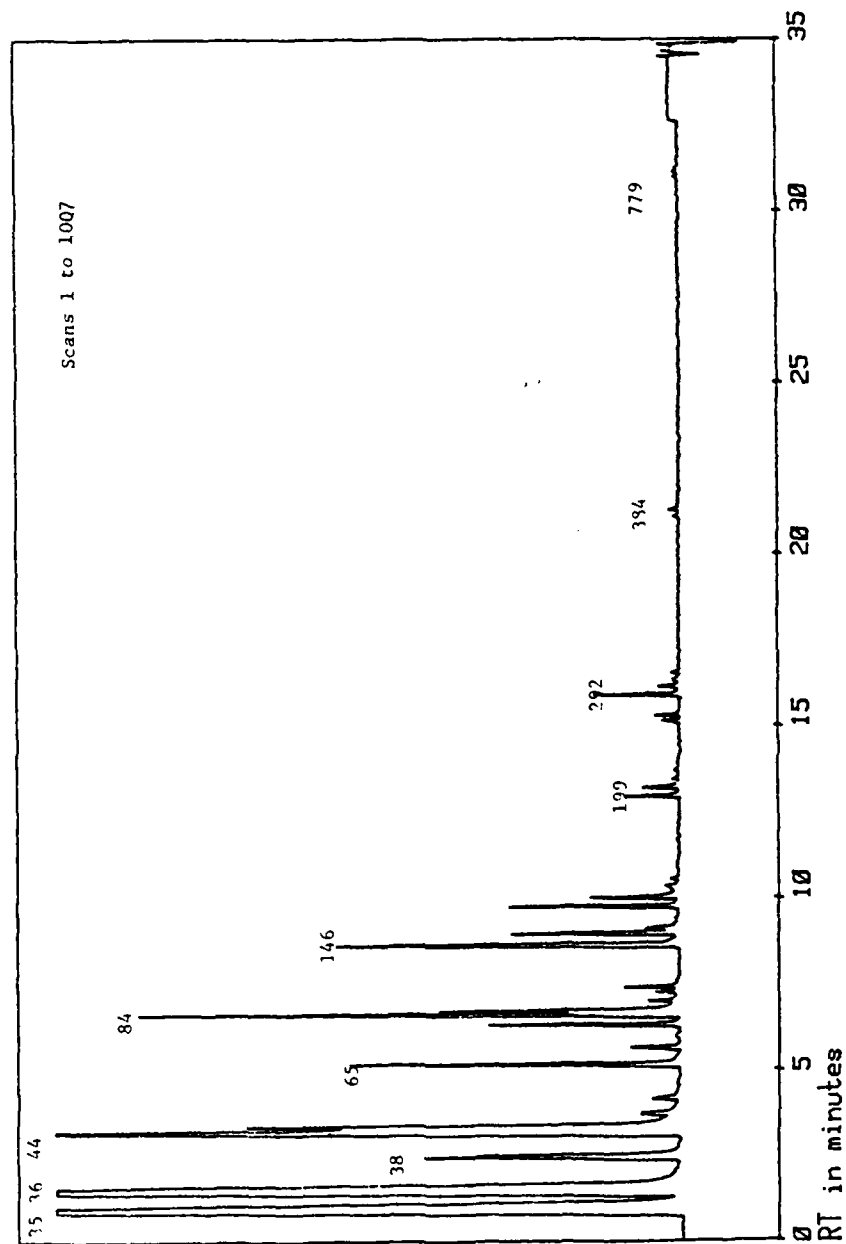


Fig 4: CHROMATOGRAM MLO 84-276, 700F, 6 HOURS HEADSPACE

Table 9 : GC/MS Analysis of Gaseous Components of MLO 84-284 Stressed at 371.1°C for 7 Hours.

<u>Scan No.</u>	<u>Component</u>	<u>% Concentration</u>
14	$\text{CH}_4$	31.48
20	$\text{C}_2\text{H}_6$	29.30
41	$\text{C}_3\text{H}_6$ and $\text{C}_3\text{H}_8$	15.15
50,52	$\text{C}_4\text{H}_8$ and $\text{C}_4\text{H}_{10}$	2.10
96,93,101,106	$\text{C}_5\text{H}_{10}$ and $\text{C}_5\text{H}_{12}$	11.99
170	$(\text{CH}_3)_2\text{SiH C}_3\text{H}_7$	0.06
188,201	$\text{C}_6\text{H}_{12}$ and $\text{C}_6\text{H}_{14}$	4.43
294,306	$\text{C}_7\text{H}_{14}$ and $\text{C}_7\text{H}_{16}$	2.11
386,394,396	$\text{C}_8\text{H}_{16}$ and $\text{C}_8\text{H}_{18}$	1.88
466	$\text{C}_9\text{H}_{18}$	0.05
537	$\text{C}_{10}\text{H}_{20}$	0.21

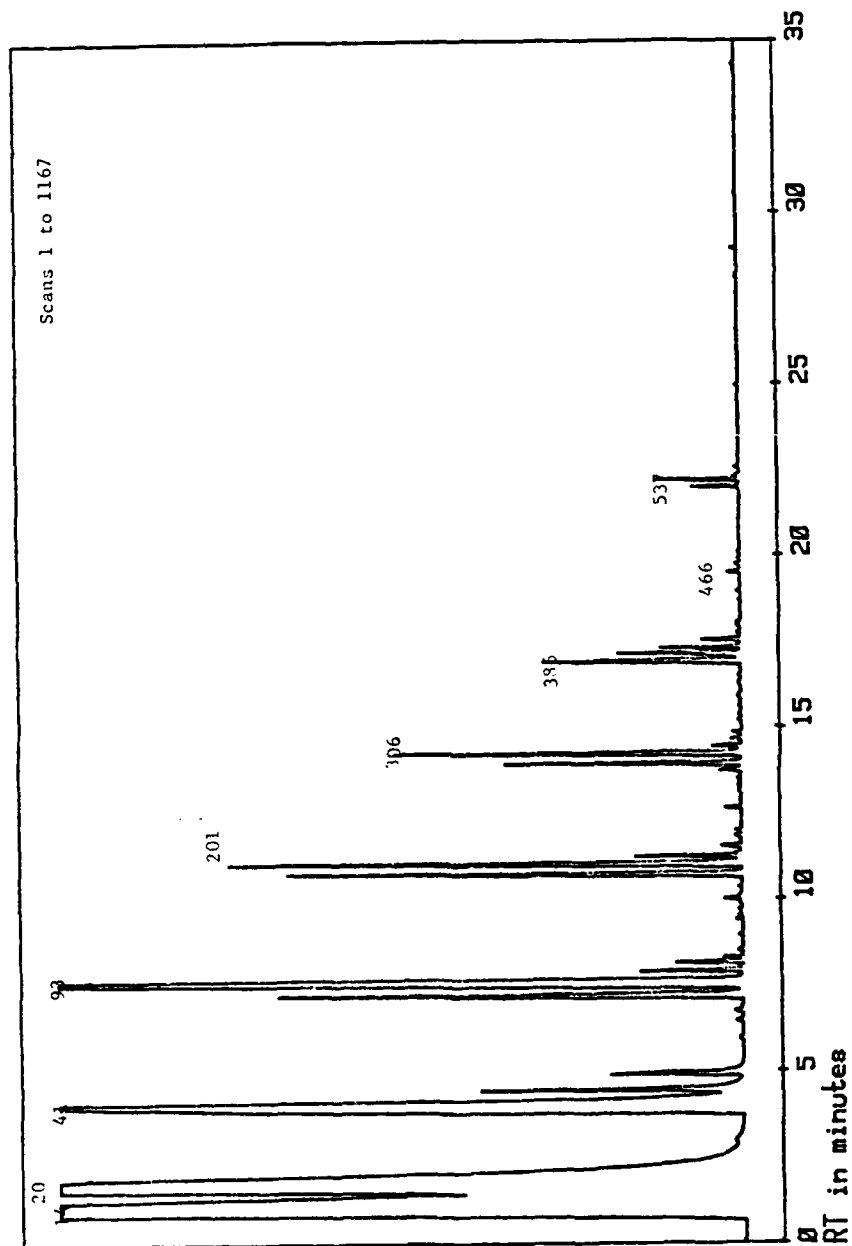


Fig. 5: CHROMATOGRAM MLO 84-284 700F 7 HOURS HEADSPACE AMPLITUDE X 2

Table 10 : GC/MS Analysis of Gaseous Components of MLO 84-285 Stressed at 371.1°C for 6 Hours.

<u>Scan No.</u>	<u>Component</u>	<u>% Concentration</u>
14	$\text{CH}_4$	20.79
20	$\text{C}_2\text{H}_6$	55.34
38	$\text{C}_3\text{H}_6$ and $\text{C}_3\text{H}_8$	8.57
48,50,55	$\text{C}_4\text{H}_{10}$ and $\text{C}_4\text{H}_{12}$	1.21
67	$(\text{CH}_3)_4\text{Si}$	0.27
82,84,90,97	$\text{C}_5\text{H}_{10}$ and $\text{C}_5\text{H}_{12}$	4.09
106	$(\text{CH}_3)_3\text{Si C}_2\text{H}_5$	1.11
185	$\text{C}_6\text{H}_{12}$	2.42
191	$\text{H Si (CH}_3)_2 \text{C}_3\text{H}_7$	0.79
199	$\text{C}_6\text{H}_{14}$	0.74
255,262	$(\text{CH}_3)_3\text{Si C}_3\text{H}_7$ and its isomer	2.32
292,303	$\text{C}_7\text{H}_{14}$ and $\text{C}_7\text{H}_{16}$	0.49
363,369	$(\text{CH}_3)_2\text{Si C}_2\text{H}_5 \text{C}_3\text{H}_7$ and its isomer	0.15
394	$\text{C}_8\text{H}_{18}$	0.08
433,437	$\text{H Si C}_2\text{H}_5(\text{C}_3\text{H}_7)_2$ and its isomer	0.15

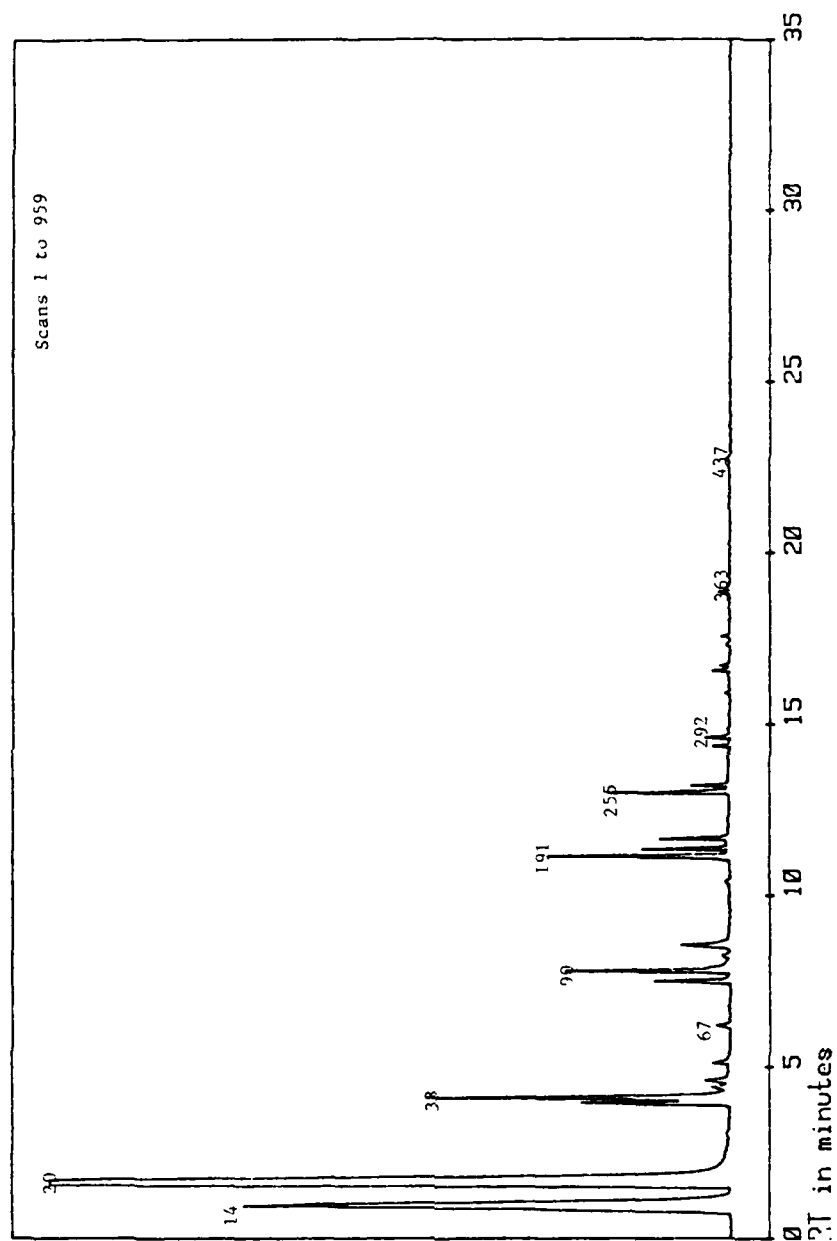


FIG. 6: CHROMATOGRAM MLO 84-285, 700F, 6HOURS HEADSPACE

Table 11: GC/MS Analysis of Liquid Components of MLO 84-193 Distillate 6  
Stressed at 371.1°C for 8 hours.

Scan No.	Component	% Concentration
48	C <sub>5</sub> H <sub>10</sub> and C <sub>5</sub> H <sub>12</sub>	0.70
52	C <sub>6</sub> H <sub>12</sub> and C <sub>6</sub> H <sub>14</sub>	0.45
62	C <sub>7</sub> H <sub>14</sub> and C <sub>7</sub> H <sub>16</sub>	1.04
79	C <sub>8</sub> H <sub>16</sub> and C <sub>8</sub> H <sub>18</sub>	0.84
102	C <sub>9</sub> H <sub>18</sub> and C <sub>9</sub> H <sub>20</sub>	0.46
137,142,146	C <sub>10</sub> H <sub>20</sub> and C <sub>10</sub> H <sub>22</sub>	2.52
183	H Si(CH <sub>3</sub> ) <sub>2</sub> C <sub>10</sub> H <sub>21</sub>	0.09
313	H Si CH <sub>3</sub> C <sub>2</sub> H <sub>5</sub> C <sub>10</sub> H <sub>21</sub>	0.10
332	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>2</sub> H <sub>5</sub> C <sub>10</sub> H <sub>21</sub>	0.20
426	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>5</sub> H <sub>11</sub> C <sub>10</sub> H <sub>21</sub>	0.11
518	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>8</sub> H <sub>17</sub> C <sub>10</sub> H <sub>21</sub>	0.14
546	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>9</sub> H <sub>19</sub> C <sub>10</sub> H <sub>21</sub>	0.11
566	H Si CH <sub>3</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.29
575	(CH <sub>3</sub> ) <sub>2</sub> Si(C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	3.13
595	CH <sub>3</sub> Si C <sub>2</sub> H <sub>5</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	2.39
618	CH <sub>3</sub> Si C <sub>3</sub> H <sub>7</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.37
638	CH <sub>3</sub> Si C <sub>4</sub> H <sub>9</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.13
659	CH <sub>3</sub> Si C <sub>5</sub> H <sub>11</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	0.99
681	CH <sub>3</sub> Si C <sub>6</sub> H <sub>13</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.13
704	CH <sub>3</sub> Si C <sub>7</sub> H <sub>15</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.24
730	CH <sub>3</sub> Si C <sub>8</sub> H <sub>17</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.64
760	CH <sub>3</sub> Si C <sub>9</sub> H <sub>19</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.51
802	CH <sub>3</sub> Si (C <sub>10</sub> H <sub>21</sub> ) <sub>3</sub>	63.91
830	CH <sub>3</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub> Si O Si <sup>+</sup> H C <sub>2</sub> H <sub>5</sub>	1.73

FIG 31 84 13:00:00  
 DATE: 275164 #1  
 CALI: CAL12 #3  
 SAMPLE: M L10010  
 COND.: GC-EL  
 RANGE: 0 1-1225 LABEL: H 0, 4.0 QUANT: H 0, 1.0 J 0 BASE: U 20, 3  
 421376

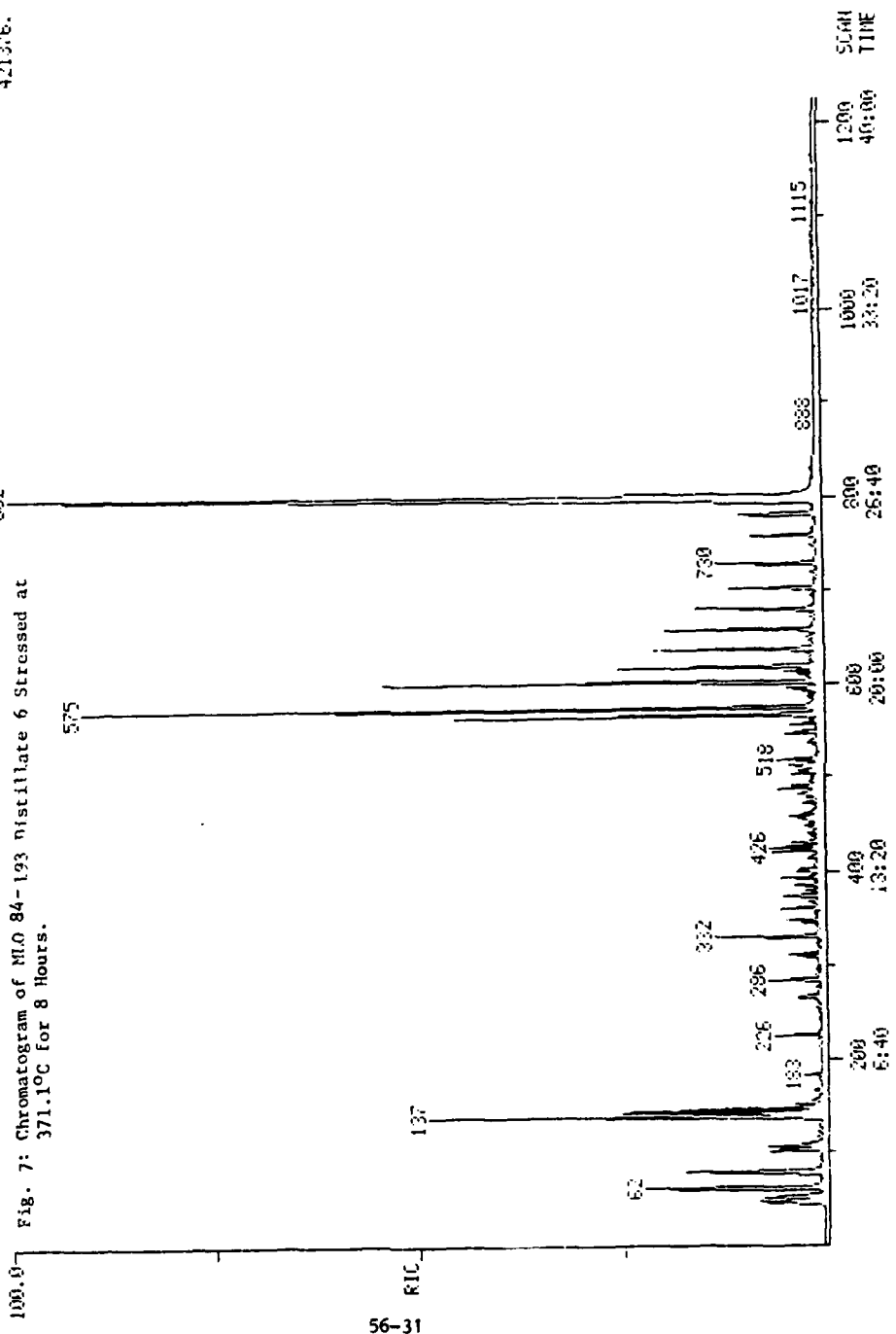


Table 12: GC/MS Analysis of Liquid Components of MLO 84-276 Stressed at 371.1°C for 8 Hours.

Scan No.	Component	% Concentration
46	C <sub>5</sub> H <sub>10</sub> and C <sub>5</sub> H <sub>12</sub>	1.27
50	C <sub>6</sub> H <sub>12</sub> and C <sub>6</sub> H <sub>14</sub>	1.22
60	C <sub>7</sub> H <sub>14</sub> and C <sub>7</sub> H <sub>16</sub>	0.48
75 77	C <sub>8</sub> H <sub>16</sub>	1.84
126	C <sub>10</sub> H <sub>22</sub>	0.17
199	(CH <sub>3</sub> ) <sub>3</sub> Si C <sub>8</sub> H <sub>17</sub>	0.20
361	H Si C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.16
374	(CH <sub>3</sub> ) <sub>2</sub> Si C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.97
411	CH <sub>3</sub> Si C <sub>2</sub> H <sub>5</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.73
433	CH <sub>3</sub> Si C <sub>3</sub> H <sub>7</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.42
450	CH <sub>3</sub> Si C <sub>4</sub> H <sub>9</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.35
455	(CH <sub>3</sub> ) <sub>2</sub> Si (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	0.57
460	CH <sub>3</sub> Si C <sub>4</sub> H <sub>9</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.31
488	CH <sub>3</sub> Si C <sub>2</sub> H <sub>5</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	2.07
504	CH <sub>3</sub> Si C <sub>3</sub> H <sub>7</sub> C <sub>4</sub> H <sub>9</sub> C <sub>8</sub> H <sub>17</sub>	0.65
516	CH <sub>3</sub> Si (C <sub>6</sub> H <sub>13</sub> ) <sub>2</sub> C <sub>8</sub> H <sub>17</sub>	0.53
522	CH <sub>3</sub> Si C <sub>4</sub> H <sub>9</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	0.27
543	CH <sub>3</sub> Si C <sub>7</sub> H <sub>15</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.38
561,574	CH <sub>3</sub> SiC <sub>6</sub> H <sub>13</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	55.94
605	CH <sub>3</sub> SiC <sub>7</sub> H <sub>15</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	0.33
622	CH <sub>3</sub> C <sub>2</sub> H <sub>5</sub> C <sub>8</sub> H <sub>17</sub> Si-O-SiCH <sub>3</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	0.86
630	CH <sub>3</sub> Si (C <sub>8</sub> H <sub>17</sub> ) <sub>3</sub>	2.20
684	CH <sub>3</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub> Si-O-SiCH <sub>3</sub> C <sub>6</sub> H <sub>13</sub> C <sub>8</sub> H <sub>17</sub>	6.84



RIC  
 07/30/84 13:32:00  
 SAMPLE: #17LIQUID  
 COND5.: GC-EI  
 RANGE: C 1.1139 LABEL: N 0, 4.0 QUANT: A 0, 1.0 J 0 BASE: U 20, 3  
 DATA: 27512 #1  
 CALI: CAL12 #3  
 SCANS 1 TO 1199

1202170.

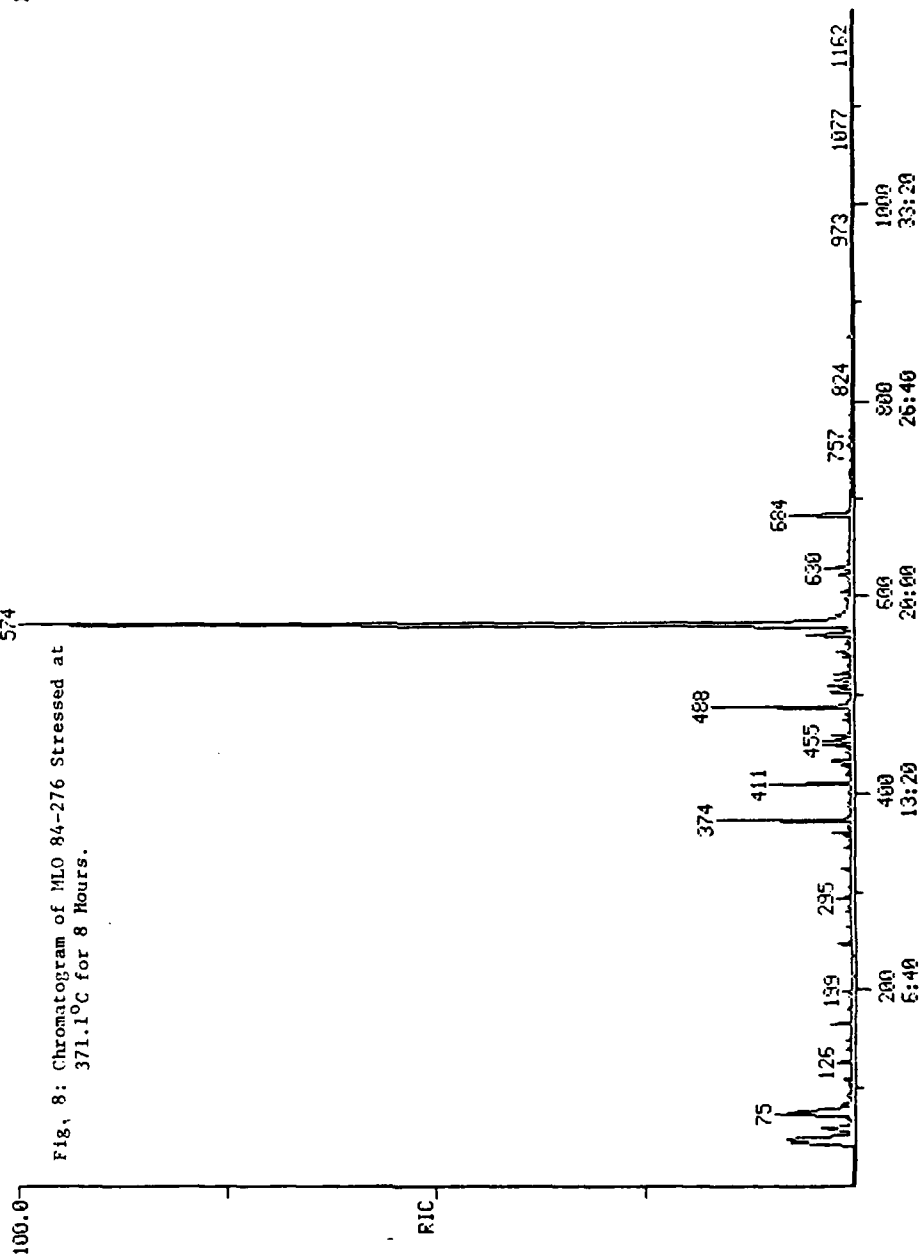


Table 13:GC/MS Analysis of Liquid Components of MLO 84-284 Stressed at 371.1°C for 8 Hours.

Scan No.	Component	% Concentration
45	C <sub>5</sub> H <sub>10</sub> and C <sub>5</sub> H <sub>12</sub>	0.67
50	C <sub>6</sub> H <sub>12</sub> and C <sub>6</sub> H <sub>14</sub>	0.48
59	C <sub>7</sub> H <sub>14</sub> and C <sub>7</sub> H <sub>16</sub>	0.68
76	C <sub>8</sub> H <sub>16</sub> and C <sub>8</sub> H <sub>18</sub>	1.43
100,104	C <sub>9</sub> H <sub>18</sub> and C <sub>9</sub> H <sub>20</sub>	0.25
135,140,144	C <sub>10</sub> H <sub>20</sub> and C <sub>10</sub> H <sub>22</sub>	1.33
249	(CH <sub>3</sub> ) <sub>2</sub> SiC <sub>2</sub> H <sub>5</sub> C <sub>8</sub> H <sub>17</sub>	0.08
295	CH <sub>3</sub> Si(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> C <sub>8</sub> H <sub>17</sub>	0.11
373	CH <sub>3</sub> Si(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> C <sub>10</sub> H <sub>21</sub>	0.12
475	H SiC <sub>2</sub> H <sub>5</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	0.74
488	CH <sub>3</sub> SiC <sub>2</sub> H <sub>5</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	1.37
518	(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> Si (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	1.22
535	C <sub>2</sub> H <sub>5</sub> Si C <sub>3</sub> H <sub>7</sub> (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub>	0.64
590	H Si C <sub>2</sub> H <sub>5</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.16
600	C <sub>2</sub> H <sub>5</sub> Si CH <sub>3</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.73
626	(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub> Si(C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	2.41
639	C <sub>2</sub> H <sub>5</sub> Si C <sub>3</sub> H <sub>7</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	0.51
653	C <sub>2</sub> H <sub>5</sub> Si (C <sub>8</sub> H <sub>17</sub> ) <sub>3</sub>	30.37
674	C <sub>2</sub> H <sub>5</sub> Si C <sub>7</sub> H <sub>15</sub> C <sub>8</sub> H <sub>17</sub> C <sub>10</sub> H <sub>21</sub>	0.80
697	C <sub>2</sub> H <sub>5</sub> Si (C <sub>8</sub> H <sub>17</sub> ) <sub>2</sub> C <sub>10</sub> H <sub>21</sub>	3.89
724	C <sub>2</sub> H <sub>5</sub> Si C <sub>7</sub> H <sub>15</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	1.18
753	C <sub>2</sub> H <sub>5</sub> Si C <sub>8</sub> H <sub>17</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	3.65
786	C <sub>2</sub> H <sub>5</sub> Si C <sub>9</sub> H <sub>19</sub> (C <sub>10</sub> H <sub>21</sub> ) <sub>2</sub>	0.59
829	C <sub>2</sub> H <sub>5</sub> Si (C <sub>10</sub> H <sub>21</sub> ) <sub>3</sub>	29.26

RIC  
 07/30/84 15:05:00  
 SAMPLE: HOLLY-LIQUID  
 CONDUS.: GC-E1  
 RANGE: G 1,1117 LABEL: N 0, 4.0 OURN: A 0, 1.0 J 0 BASE: U 20, 3  
 DATA: 27514 #1  
 CALI: CAL12 #3  
 SCANS 1 TO 1117

531872.

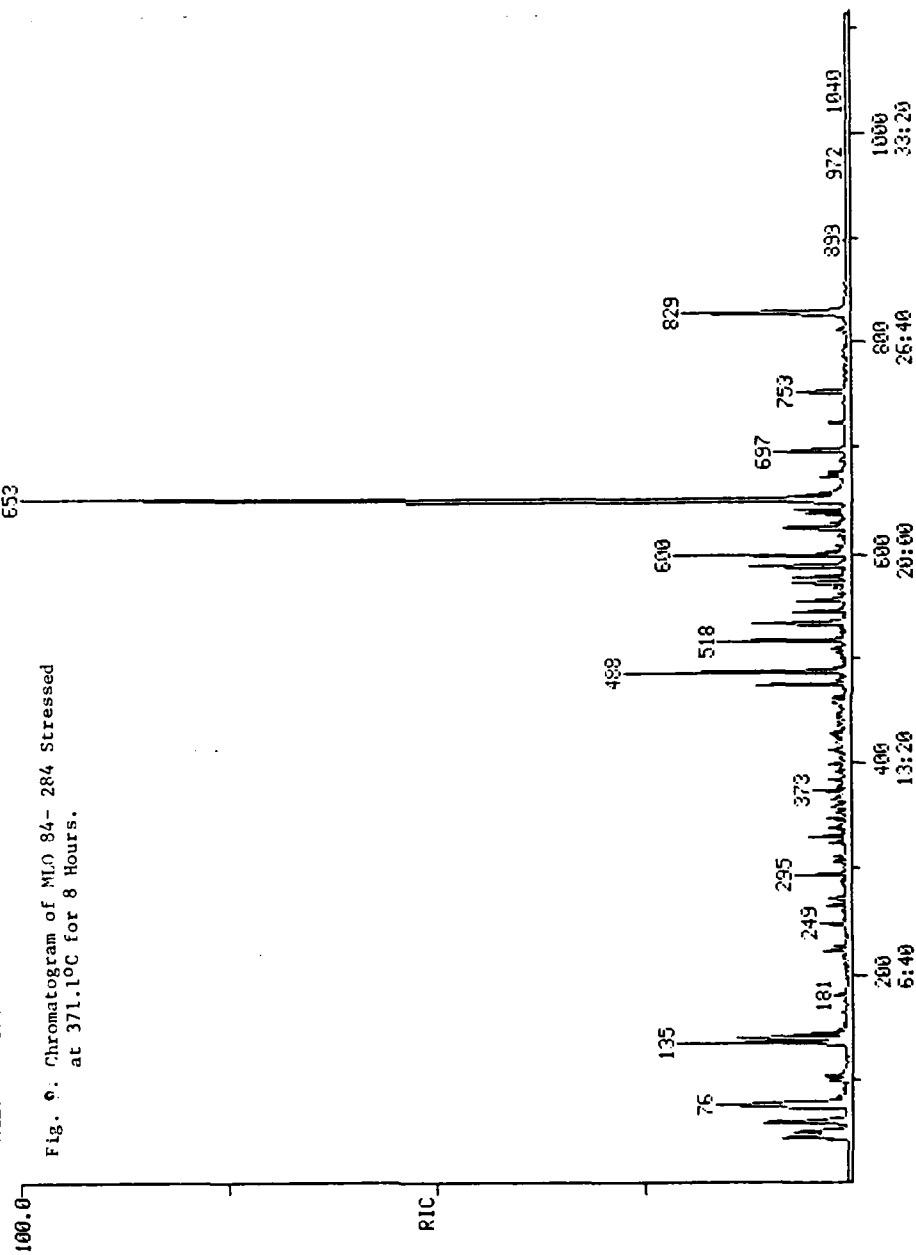


Table 14 GC/MS Analysis of Liquid Components of MLO 84-285 Stressed at 371.1°C for 8 Hours.

Scan No.	Component	% Concentration
46	$C_4H_{10}$ , $C_5H_{10}$ , $C_5H_{12}$ , $C_6H_{12}$	0.41
50	H Si $(CH_3)_2$ $C_3H_7$	0.41
54	$(CH_3)_3$ Si $C_3H_7$	0.58
59	$C_7H_{16}$	0.20
70	H Si $C_3H_7(C_2H_5)_2$ or $(CH_3)_2$ Si $C_2H_5C_3H_7$	0.37
76	$C_8H_{18}$	0.22
89	$(CH_3)_2$ Si $(C_3H_7)_2$	0.71
103	$C_9H_{20}$	0.21
139	$C_{10}H_{22}$	0.23
180	$C_{11}H_{24}$	0.45
224	$C_{12}H_{26}$	0.30
240	$(CH_3)_2$ Si $C_3H_7$ $C_7H_{15}$	0.21
282	$(CH_3)_2$ Si $C_3H_7$ $C_8H_{17}$	0.24
304	$C_{14}H_{28}$	0.68
309	$C_{14}H_{28}$ and $C_{14}H_{30}$	0.98
322	$(CH_3)_2$ Si $C_3H_7C_9H_{19}$	0.24
361	$(CH_3)_2$ Si $C_3H_7C_{10}H_{21}$	0.24
418	$(CH_3)_2$ Si $C_2H_5$ $C_{13}H_{27}$	0.27
454	$C_{18}H_{38}$	0.35
466	$(CH_3)_2$ Si $C_3H_7C_{13}H_{27}$	0.43
510	$(CH_3)_2$ Si $C_3H_7$ $C_{14}H_{29}$	78.71
726	$(CH_3)_2$ Si $C_{12}H_{25}C_{14}H_{29}$	0.21
794	$(CH_3)_2$ Si $(C_{14}H_{29})_2$	0.91
837	H $CH_3C_{14}H_{29}$ Si-O-Si <sup>+</sup> HC <sub>3</sub> H <sub>7</sub>	0.16

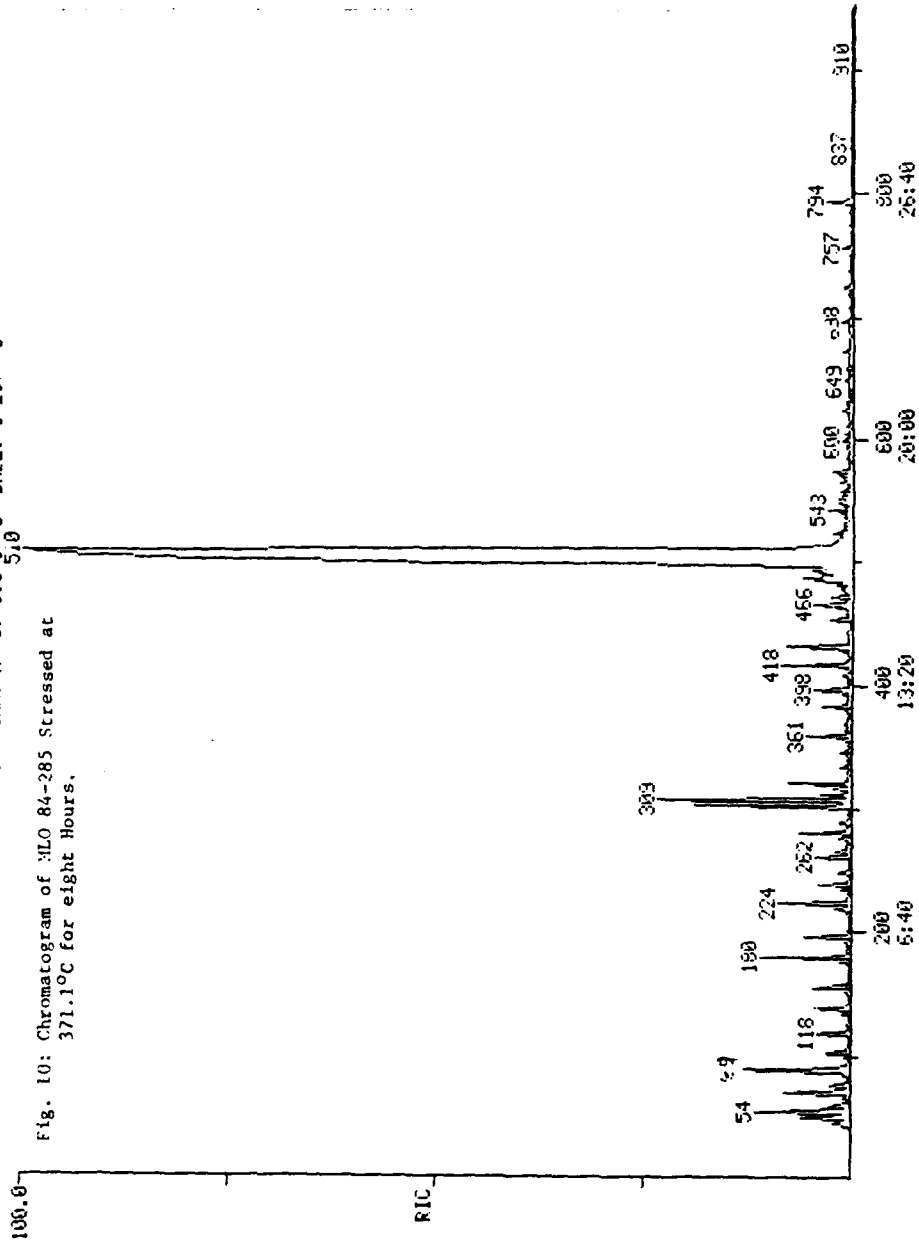
PIC  
08/01/84 13:10:00  
SAMPLE: #4 LIQUID  
COMPOS.: GC-EI  
RANGE: 0 1.954

DATA: 27510 #1  
CALI: CALI3 #3

SCHMIS 1 TO 354

LABEL: N 0, 4.0 UNK: A 0, 1.0 J 0 BASE: U 20, 3

1533710.



#### REFERENCES

1. Snyder, C. E. Jr., Gschwender, L.J., and Tamborski C., 'Synthesis and Characterization of Silahydrocarbons - A class of Thermally Stable Wide-liquid Range Functional Fluids", ASLE Trans. 25, 299-303(1981).
2. Snyder, C.E. Jr., Tamborski, C. Gschwender, L.J., and Chen, G.J., "Development of High Temperature (  $-40^{\circ}\text{C}$  to  $289^{\circ}\text{C}$ ) Hydraulic Fluids for Advanced Aerospace Applications", ASLE Preprint No. 80-LC-IC-1.
3. Snyder, C.E. Jr., and Gschwender, L.J., "Development of a Non-flammable Hydraulic Fluid for Aerospace Applications over a  $-54^{\circ}\text{C}$  to  $135^{\circ}\text{C}$  Temperature range", Lubr. Engr. 36, 458 (1980).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

CALCULATIONS OF ELECTRON SPIN RESONANCE COUPLING CONSTANTS

Prepared by: Dr. Hendrik F. Hamka  
Academic Rank: Professor  
Department and University: Department of Chemistry  
University of Pennsylvania, Philadelphia, PA  
Research Location: Frank J. Seiler Research Laboratory  
USAF Academy, Colorado Springs, CO 80840  
USAF Research Contact: Major Larry P. Davis  
Date: 23 August 1984  
Contract No: F49620-82-C-0035

Calculations of Electron Spin  
Resonance Coupling Constants

by

Hendrik F. Hamerka

Abstract

One of the current research areas of the Frank J. Seiler Research Laboratory is the thermal decomposition of energetic materials. A major part of the experimental information of these types of chemical reactions is derived from electron spin resonance (ESR) experiments. In order to interpret the experimental data, it is useful to have theoretical predictions of the ESR coupling constants of various possible reaction intermediates. We feel that the most effective procedure for calculating ESR coupling constants is the use of the unrestricted Hartree-Fock method (UHF). We first verified the accuracy of UHF results by calculating the ESR coupling constants of some typical organic radicals. The agreement between theory and experiment was generally poor, and we attributed the poor agreement to spin contamination of the UHF wave function. We derived a simple procedure for removing the spin contamination from the UHF function, and we found that the corrected results agree satisfactorily with experiment. We plan to utilize this procedure for theoretical predictions on organic radicals that are relevant to the thermal decomposition of energetic materials.



#### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend an interesting and productive summer at the Frank J. Seiler Research Laboratory.

In particular, he wishes to thank Dr. John S. Wilkes, Maj Larry P. Davis, Lt Jon T. Swanson, and Dr. Almon G. Turner for suggesting this area of research and for many helpful discussions. He also wishes to express his appreciation for the excellent computational, secretarial, and technical service of the Frank J. Seiler Research Laboratory.

## I. Introduction

One of the current areas of research within the Frank J. Seiler Research Laboratory is the thermal decomposition of energetic materials. A typical example is the thermal decomposition of trinitrotoluene (TNT). The thermal decomposition of TNT involves the formation of various unknown reaction intermediates. A major part of the experimental information of the thermal decomposition reaction is derived from electron spin resonance (ESR) experiments. In order to interpret the ESR data it is desirable to know the values of the ESR proton and carbon coupling constants of various organic free radicals that are possible intermediates in the thermal decomposition reaction.

Obviously it would be helpful if these various coupling constants could be predicted theoretically. The ESR coupling constant between a particular nucleus and the unpaired electron is derived by calculating the total electron density of all alpha electrons and the total electron density of all beta electrons at the position of the nucleus. The coupling constant is proportional to the difference between the two electron densities. A calculation of the ESR coupling constant requires first a knowledge of the molecular wave function and second, a suitable computational procedure for deriving the coupling constant from the wave function.

At the Frank J. Seiler Research Laboratory, the computing facility has the capability for deriving molecular wave functions by means of ab initio methods via the Gaussian 82 program package and by means of various semi-empirical methods. In all these methods, the wave functions are derived by means of the Hartree-Fock approximation, that is, they are obtained as anti-symmetrized products of molecular orbitals.

There are two different procedures for calculating ESR coupling constants by means of Hartree-Fock methods. In the first procedure, the restricted Hartree-Fock method (RHF), we impose the requirements that the alpha and beta orbitals are partitioned into pairs with identical orbital parts. In addition, there is one electron in a singly occupied orbital with either alpha or beta spin. The RHF method by itself does not lead to reliable theoretical results for ESR coupling constants, but if the wave function is corrected by adding configuration interaction, it may yield reliable theoretical values.

The second procedure is the unrestricted Hartree-Fock (UHF) method. Here the alpha and beta orbitals are allowed to be different. The result is an anti-symmetrized UHF function, but in general this function is not an eigenfunction of the total spin operator  $S^2$ . In other words, a UHF function does not belong to a proper spin state. Since the correct wave function should belong to a proper spin state with total spin  $S = (1/2)$ , it is not at all clear why an UHF function should produce reliable values for ESR coupling constants.

Even though I had no previous experience in the evaluation of ESR coupling constants, I have worked extensively in the general theoretical area of magnetic resonance and of the interaction between molecules and magnetic fields. The Frank J. Seiler Research Laboratory has good computational facilities for the derivation of molecular wave function, and the research staff is interested in deriving procedures for computing ESR coupling constants. I was attracted by the good quality of the computational facilities and by the challenge of the problem of deriving simple and accurate procedures for calculating ESR coupling constants.

## II. Objectives

Our objective is the derivation of simple and accurate procedures for calculating ESR coupling constants which are compatible with the computing facilities of the Frank J. Seiler Research Laboratory and which are applicable to the organic radicals that are of interest in the thermal decomposition of energetic materials research.

Our initial plan was to perform a series of calculations on a variety of small and intermediate-size molecules by means of both ab initio and semi-empirical methods. The ab initio calculations were to be performed by means of the Gaussian 82 program package.

We felt that the unrestricted Hartree-Fock (UHF) method constituted the most practical and effective approach to the calculation of ESR coupling constants. In order to test the accuracy of the method, we performed some calculations on small organic radicals for which experimental data were available. We found the agreement between these calculations and experiments to be generally poor. In some isolated instances, the agreement between

theory and experiment was good, but in other cases the discrepancies were very large, and there did not seem to be any clear correlation between experimental and theoretical ESR coupling constants. We felt therefore, that it was necessary to analyze the theoretical procedures and to derive a method that led to reliable values before we should consider larger systems. We therefore changed our initial research objective of comparing ab initio and semi-empirical theoretical results for large molecules. Instead we set an alternate goal of deriving a reliable theoretical procedure that leads to reasonable agreement with experiment for some smaller organic radicals.

We have some indications that the latter goal is not as easy to achieve as it seems. Chipman<sup>1</sup> published some very accurate calculations on ESR coupling constants of small molecules, in particular, on the methyl radical. Considering the quality of the calculations, the results are disappointing; discrepancies of a factor two or more are not unusual. We felt therefore, that a careful analysis of the theoretical procedures is necessary in order to obtain accurate theoretical results from Hartree-Fock calculations.

### III. Theoretical analysis

The most convenient procedure for evaluating ESR coupling constants is by means of the unrestricted Hartree-Fock (UHF) method. Unfortunately, the results that are obtained by means of this method are usually poor. We mentioned already that the molecular wave function which is the result of a UHF calculation is usually not an eigenfunction of the operator  $S^2$  where  $S$  is the total spin of the molecule. Therefore, a UHF wave function is not realistic from a physical point of view, and it should be corrected in order to make it acceptable from a physical point of view.

The ground state of an organic radical should have  $S = 1/2$ , spin multiplicity 2 but a UHF wave function  $\Psi_0$  is a linear combination of various spin eigen functions,

$$\Psi_0 = {}^2\Psi_0 + {}^4\Psi_0 + {}^6\Psi_0 + \text{-----} \quad (1)$$

The higher order spin functions are called the spin contamination of the wave function because they should not be there, and they should be

removed in order to obtain correct theoretical results from the function  ${}^2\Psi_0$ .

Various procedures for removing the unwanted spin functions from a UHF function have been proposed.<sup>2,3,4,5</sup> These procedures are basically correct, but they are rather complicated. We were interested in deriving a simpler procedure which is more compatible with the computer programs that are available.

From the various calculations that we performed, we noted that most UHF wave functions exhibit certain features that may be utilized as a starting point of a theoretical analysis. We mentioned that in a RHF calculation we impose the restriction that the alpha and beta orbitals form pairs with identical orbital functions. In addition, there is one unpaired electron which we assign to an alpha orbital. In a UHF calculation, this constraint is removed, nevertheless, we usually observe pairs of alpha and beta orbitals with different but similar orbital parts. Therefore, we assume that the orbitals of such a pair may be written as

$$\begin{aligned}\varphi_i^\alpha(i)\alpha(i) &= [\chi_i(i) + \Delta_i(i)]\alpha(i) \\ \varphi_i^\beta(N+i)\beta(N+i) &= [\chi_i(N+i) - \Delta_i(N+i)]\beta(N+i)\end{aligned}\quad (2)$$

We assume now that the differences  $\Delta_i$  between the orbitals are much smaller than the orbitals  $\chi_i$  so that it is permissible to expand the UHF wave function  $\Psi_0$  as a power series in the quantities  $\Delta_i$ . We also assume that all terms that are quadratic or high order in  $\Delta_i$  may be neglected. It is then possible to determine the doublet spin function  ${}^2\Psi_0$  of Eq. (1) exactly.

We showed also that the removal of the unwanted spin functions in the above described model leads to a particularly simple procedure for correcting the UHF results for the ESR coupling constants. The results are corrected by multiplying all contributions from the paired alpha and beta orbitals by a factor one-third and by multiplying the contribution from the unpaired orbital by a factor unity.

#### IV. Calculations

We verified the accuracy of our theoretical analysis by performing calculations on four organic radicals by means of the Gaussian 82 program package. As our samples, we selected methyl, allyl, ethyl, and vinyl. In the first two cases, methyl and allyl, the unpaired orbital does not contribute to the ESR coupling constants, and the calculations become quite simple. All we had to do was to correct the previous UHF results by multiplying them by a factor one-third. In the third and fourth cases, the singly occupied orbital contributes to the ESR coupling constants. Here we had to evaluate the contribution by hand in order to derive the final corrected results. However, the procedure was reasonably straightforward, and we feel that the procedure may be incorporated into the various computer programs without much difficulty.

We described the various calculations and results in a manuscript that we submitted for publication in the Journal of the American Chemical Society. We enclosed a preprint of this ms. This preprint contains all details of our theoretical analysis and the results of our calculations. It should be noted that the agreement between theory and experiment is quite satisfactory.

#### V. Recommendations

We feel that our theoretical analysis and its conclusions constitutes a major contribution to the theory of ESR coupling constants. The advantage of our procedure is that it may be coupled with available computer programs, in particular, the G82 program package. By means of a few extremely simple additional steps, we may obtain theoretical results with accuracies that are comparable to the accuracies of results that are derived by much more complicated theoretical procedures.

We should note that it is possible to obtain reasonably good ESR coupling constants by alternative procedures, for instance, from RHF calculations supplemented by configuration interaction or from UHF calculations where the spin contamination is removed by means of projection operators. However, our method is simpler and this means that it is applicable to larger systems than the other methods are.

The computational facilities at the Frank J. Seiler Research Laboratory are excellent, but there are limitations to the size of molecules that may be tackled. We feel that ab initio calculations with the G82 program package cannot be performed on molecules that are larger than mono nitro toluene, otherwise the amounts of computer time involved becomes excessive. It is possible to perform calculations on larger molecules by means of semi-empirical procedures but in those cases it is hard to predict the accuracy of the results.

We feel that our work has supplied us with the tools to tackle our initial research goal. We would like to perform calculations on some medium size organic radicals such as phenyl, benzyl, nitro propyl, etc. by means of the ab initio G82 procedures and compare those results with experimental results and with semi-empirical results. Ultimately, we wish to derive theoretical procedures that may be reduced to a form that may conveniently be used by experimentalists. We hope that the experimentalists may be supplied with theoretical procedures that are easy to use and that are useful for interpreting experimental data and for elucidating reaction mechanisms for the thermal decomposition of energetic materials.

We plan to submit a research proposal that will enable us to pursue the research objectives that we have described above.

### References

1. D. M. Chipman, J. Chem Phys. 71, 761 (1979); 70, 3112, 4725 (1983).
2. P. O. Löwdin, Adv. in Chem. Phys. 2, 207 (1958).
3. J. E. Harriman, J. Chem. Phys. 40, 2827 (1964).
4. T. Amos and L. C. Snyder, J. Chem. Phys. 41, 1773 (1964).
5. D. H. Phillips and J. C. Schug, J. Chem. Phys. 61, 1031 (1974).



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Effects of Pyridostigmine on Performance of Mission-ready Pilots in the

OT Simulation Facility

Prepared by:	Arthur E. Harriman
Academic Rank:	Research Professor
Department and University:	Department of Psychology, Oklahoma State University
Research Location:	Human Resources Laboratory, Williams Air Force Base
USAF Research	Mr. Robert Woodruff
Date:	August 20, 1984
Contract No:	F49620-82-C-0035

Effects of Pyridostigmine on Performance of Mission-ready Pilots in the  
OT Simulation Facility

by  
Arthur E. Harriman

ABSTRACT

A study was designed for the purpose of studying the effects of pyridostigmine on pilot performance in the AFHRL/OT flight simulation facility. In this way, usefulness of the drug as an agent of chemical defense for USAF aircrews against nerve agents of the organophosphorus anti-cholinesterase series may be further assessed. The study was planned so that the subjects, who are mission-ready pilots, will be tested in the flight simulator during the course of a double-blind procedure in which the pilots are or are not under the influence of the drug while wearing or not wearing protective gear. Permutations of the procedure will be counter-balanced among subjects. Both behavioral and physiological measures will comprise the dependent variables. The results will be treated by analysis of variance for an experiment of the treatments-by-treatments-by-subjects design. At all times, the subjects will be monitored by a licensed physician in the possibility that there are side-effects from this medically useful and well-tolerated drug.

#### ACKNOWLEDGEMENTS

I wish to express my deep sense of gratitude to the Air Force Systems Command, Air Force Office of Scientific Research, for its sponsorship of the USAF-SCEEE Faculty Research Program in which I participated during the summer of 1984. The program was, in my opinion, highly influential in broadening my understanding of research in the areas of psychopharmacology, psychotoxicology, and psychophysiology. My background in physiological psychology now seems to have been narrower in scope than I had viewed it prior to my period of work at Williams Air Force Base. As it stands now, therefore, I think that I will be improved in my teaching and research when I return to my home institution.

Additionally, I am grateful to the Human Resources Laboratory personnel who worked with me on various phases and aspects of the summer project. In this regard, I am particularly grateful to Mr. Robert Woodruff and to Mr. Robert Bunker who were unfailingly pleasant, always helpful, and admirably cooperative and responsive in their collegial association with me. During what was in retrospect a strongly productive and professionally enriching period of work, both of these individuals stand out in memory as patient, understanding, and talented associates.

No statement in acknowledgement of my gratitude to the AFOSR and to the personnel who implemented the USAF-SCEEE program in 1984 should end without an indication of the deep appreciation I have for the work of Dr. Warren D. Peele. As Managing Director of the USAF-SCEEE program, Dr. Peele clarified every step of action to be undertaken by the participants, responded graciously to inquiries, and proved extraordinary in the degree of conscientiousness and efficiency with which he managed direction of the program. I am pleased to learn from fellow 1984 members of the program that this high opinion of Dr. Peele is strongly shared.

## I. INTRODUCTION:

My teaching and research interests, since graduate school, have been in the area of physiological psychology, and a person with this research background was needed in the Human Resources Laboratory of Williams AFB for the task described on the accompanying pages. Within the physiological psychology area, I have been using the Mongolian gerbil in recent years as an animal model for extrapolations to human behavior. The gerbil has proved a useful model, in my opinion, for the study of idiopathic epilepsy in humans because, like these human epileptics, some proportion of gerbils seize 'spontaneously'. Several published papers report support for an inference that magnesium metabolism is at fault in the susceptible gerbils and relate this conclusion to scattered, nonrecent observations on humans contained in the medical literature. The other work has used the gerbil as a model for osmoregulatory and thermoregulatory activities in humans. In the work, analogies were attempted between coat densities and air pockets within hair layers, on one hand, and nude versus clothed skin of human, on the other, in conditions of restricted water availability and varied temperatures. In these, and similar, studies, there was considerable need for good electrophysiological monitoring and recording equipment.

Additionally, I have done some research in the area of psychopharmacology which had to do with behavioral and neurological routes of effect of convulsant and "frightening" agents upon avian and small mammal predation on crops. This line of work, like the more extensive interest referred to above, proved to be very helpful in defining the character of my contribution to the project.

The nature of this project is one requiring contributions by persons with established skills in a variety of disciplines. Several individuals in the Human Factors Laboratory were fully conversant with the specific elements for much of the project as it began to develop, such as the measurement of psychomotor skills and awareness of the visual perceptual factors involved. Nonetheless, there was the perception that the developing project fell partly outside the scope of ongoing activities in the Laboratory and that someone versed in physiological psychology, with awareness of methods used in psychopharmacology, might be helpful in planning work.

When I put in my application to the SCEEE in the hope of getting a 1984 research appointment, I was interested in going to Williams AFB because several of the strong research programs described in the SCEEE brochure paralleled my physiological interests. These programs included

efforts having to do with the human visual system and its functions in simulated military situations. Once in contact with research personnel in the Laboratory, it became possible for me to focus on a specific perceptual/motor skills problem which requires contributions from psychophysiology and psychopharmacology in addition to inputs, as mentioned, from the regular staff members.

## II. OBJECTIVES OF THE RESEARCH EFFORT:

At the outset of the summer research effort, my USAF colleague graciously cooperated with me in developing an explicit list of research objectives for the summer period. The identified objectives were as follows:

1. to review the literature on pyridostigmine as a member of the anticholinesterase series of drugs and toxins.
2. to explore the possible effects of pyridostigmine on human performance in cognitive and motor skills functions.
3. to explore the possible effects of pyridostigmine on the peripheral nervous system - both the autonomic and the somatic divisions.
4. to work with a member of the Laboratory in developing a protocol for evaluating effects on pyridostigmine on human performance in a flight simulator device.
5. to devise ways to measure most appropriately the effects of pyridostigmine on the autonomic nervous system.
6. to organize all of the above information into an acceptable research protocol which would then be suitable for implementation by the members of the Laboratory.

These objectives were not, as it turned out, approached in fully sequential fashion, though all of them were dealt with in the course of the summer period. Further, certain of the objectives were greatly expanded, while, because of limited information in the journal literature, other objectives were met in more limited manner. Nonetheless, thanks to the help of my research colleague, the research goals, which were blocked out near the outset of the work, remained viable throughout the period of the summer project. There was, however an addition to the goals. On the basis of work done in connection with objective #5 (above), it was suggested to my research colleague and myself that another objective be framed. This objective was as follows:

7. to evaluate telemetering devices versus hard-wired apparatus for valid monitoring of ongoing physiological processes in active, clothed human subjects.

### III. REVIEW OF THE JOURNAL LITERATURE ON PYRIDOSTIGMINE:

- a. Approach. I learned from my prospective research colleague during the course of my pre-summer site visit that there was considerable interest in the study of how pyridostigmine could serve as a pretreatment drug in the protection of aircrews against exposure to nerve agents. During the month which preceded my arrival at the Laboratory, I reviewed the journal literature on the subject in the library of my University. After arrival at the Laboratory, I made further use of the library at the nearest university in the same connection.
- b. Summary of the results achieved in the literature review. Chemicals in the carbamyl ester group, which functionally may be included in the "reversible" anticholinesterase (anti-ChE) class of compounds, counteract effects of toxicants in the organophosphorus, "irreversible" anti-ChE group of compounds. The carbamyl ester compound, pyridostigmine, has been recommended for investigation as a pre-exposure drug which shows promise of protecting personnel against organophosphorus nerve agents, such as tabun, sarin, and soman. The latter nerve agents have been available for several decades and apparently have been used against military and civilian populations in the contemporary period by Soviet-supplied military forces. Presumably, pyridostigmine has been selected for investigation by the USAF because it has relatively milder side effects and has longer duration of action than do the other carbamyl ester compounds.

The underlying thesis for use of pyridostigmine as a protection against subsequent exposure to organophosphorus anti-ChE compounds is that this drug, for the reasons cited, is successful in acting competitively as an enzyme inhibitor by "reversibly" occupying postsynaptic receptor sites where the neurotransmitter is acetylcholine (ACh). In consequence of this occupying action, formation of organophosphate-esterase complexes of enduring ("irreversible") character does not take place (for the time of their action) on the

postsynaptic membrane at cholinergic synapses following exposure of an organism to organophosphorus compounds.

Nonetheless, pyridostigmine, like other anti-ChE compounds or cholinomimetic agents, causes ACh to remain intact over a longer time span, thereby accumulating at cholinergic sites in nerve networks which are widely distributed in the central nervous system and the autonomic nervous system and which innervate the peripheral somatic (skeletal muscle) system. The early effect of the build-up in ACh is excitatory, but, with continuing increase in the intact ACh at cholinergic synapses, the direction is toward more massive depolarizing plus hyperpolarizing effects on the subsynaptic membrane that eventuates in cessation of nerve impulses.

In the foregoing respects, the "reversible" anti-ChE agents, such as pyridostigmine, differ in degree of severity of effect (quantitatively), rather than in kind, from the "irreversible" anti-ChE toxins. More closely looked at, all anti-ChE agents have their effect on cholinergic receptors which belong to one or the other of two categories: the muscarinic and the nicotinic. Cholinergic receptors of the nicotinic type are found in the spinal cord, within the autonomic ganglia, and at neuromuscular junctions of the peripheral somatic nervous system. All autonomic cholinergic effectors cells, however, are muscarinic. Both types of receptors are densely present in the brain.

#### IV. POSSIBLE EFFECTS OF PYRIDOSTIGMINE ON HUMAN PERFORMANCE:

- a. Approach. The objective was approached by means of searches into the journal literature in psychopharmacology and psychotoxicology.
- b. Summary of results achieved in the literature review. Searches conducted in the science libraries at two universities revealed no indication of published studies having to do with dosage-effects of pyridostigmine on higher functions in human or infrahuman organisms. Nonetheless, clinical reports with human subjects showed that tolerances for the drug vary considerably within the population. Further one report indicated that one person in 1,250 persons of European descent shows extreme intolerance to the drug.

The qualitative features of pyridostigmine effects on neurological and behavioral functions can be dramatic (dysarthria, dysphonia, dysphagia, dysdipsia, and, possibly, convulsions terminating in coma). These outcomes suggested that conventional medical opinion, which holds that moderate dosages of the drug do not cross-the blood brain barrier, may be in error. Tests on learning and memory in animals given different dosages of pyridostigmine were recommended in consequence of the literature review.

V. EFFECTS OF PYRIDOSTIGMINE ON THE PERIPHERAL NERVOUS SYSTEM:

- a. Approach. Considerable information has accumulated concerning the effects of anti-ChE drugs upon the action of the peripheral nervous system. There are, however, distinct differences in the effects produced by different members of this class of chemicals. Determination of the manner in which pyridostigmine affects the human nervous system was needed so that there could be some understanding of the physiological reactions to be expected among the subjects in the study of how this drug influences pilot performance in operation of a flight simulator. Also, the information was needed in order to devise ways to detect any developing adverse reactions so that medical personnel could be alerted to the need for clinical counter-measures. Background information for implementing these aims was obtained through review of journals and monographs in the field of psychopharmacology and examination of the medical literature on myasthenia gravis.
- b. Results. At therapeutic dosages, which range in divided administrations up to 1,500 mg/day, (mean dosage = 600/mg day), pyridostigmine is well tolerated by myasthenia gravis patients. The dosages must be individualized, however; and normal young adult subjects would be expected to show side-effects to dosages at the upper end of the therapeutic range. Lower therapeutic dosages are much better tolerated by normal subjects. Nonetheless, individuals who are cholinesterase variants, and quite likely to be entirely unaware of their anomalous metabolism, may show severe reactions to low or average dosages.

The potentiation of ACh activity by pyridostigmine may cause excessive muscarinic effects and nicotinic effects.



Adverse reactions to muscarinic stimulation include miosis, lacrimation, nasal discharge, vomiting, diarrhea, sweating, increased micturation, and bradycardia. Increase in activity at nicotinic synapses may trigger fasciculation, muscular weakness, respiratory difficulties, elevated blood pressure, and tachycardia.

Apparently, the first sign of impending toxicity is the onset of muscular weakness. Therefore, it was recommended that medical monitoring of the subjects in the study be supported by use of a repeated test for skeletal muscle strength during such periods as the subjects are under the influence of pyridostigmine. The particular test suggested was a measure of hand strength to be tested with a commercial hand dynamometer at 30-min intervals.

VI. DEVELOPING A PROCEDURE FOR EVALUATING HOW PYRIDOSTIGMINE AFFECTS PERFORMANCE IN A FLIGHT SIMULATOR:

- a. Approach. There were two different approaches which were explored in connection with the aim stated above. First, there was need to become familiar with the use of the flight simulator as a device for the study of behavior. Second, an experimental procedure had to be designed in which pyridostigmine effects on pilot performance in the simulator could be validly determined.
- b. Results. I was able to gain some familiarity with the nature and the use of the flight simulation facility in this Laboratory through the help of my research colleague. The colleague conducted a detailed investigation over a period of about three weeks in which experienced pilots were tested in a flight simulator. I was able to monitor the briefings given to the subjects, the behavior of the pilots in the simulator, and the debriefings. The second mode of approach to the aim required design of an appropriate and useful experimental procedure for the proposed investigation. In this effort, a research protocol was organized which, in outline, involved a treatments-by-treatments-by-subjects procedure. In this experimental arrangement, it is possible to test subjects under conditions in which the independent variable (pyridostigmine) is administered or not administered. Also, for each condition of the independent variable, treatment permutations are possible. For the

present study, there were two such permutations - protective gear worn versus protective gear not worn. Further, the procedure of the work was arranged so that both factors as well as both treatments within factors were counter-balanced in order of use with the different subjects. Also, a double-blind procedure was described for administration of the pyridostigmine or, in its place, a placebo.

VII. INTEGRATION OF THE WORK PERFORMED IN CONNECTION WITH OBJECTIVES I THROUGH V INTO A RESEARCH METHODOLOGY WHICH WOULD BE SUITABLE FOR IMPLEMENTATION LATER BY MEMBERS OF THE HUMAN RESOURCES LABORATORY:

- a. Approach. During the final portion of the summer period, my research colleague and I organized the results of the five lines of endeavor, described above, into a detailed statement of the logic, design, method, and implications (in light of the results, whatever turn these took) of the investigation into whether pyridostigmine could act as a pretreatment drug in the protection of pilots who were exposed to organophosphorus nerve agents. Critical to the question of whether pyridostigmine should be used as such a drug is confirmation that the drug does not adversely affect complex performance. The experimental plan which resulted from this collaboration was closely guided by an appreciation of the latter requirement.
- b. Results. The nature of the research project, as it developed from the collaboration between the research colleague and myself is fully described in a lengthy presentation. This presentation ultimately may be forwarded to Brooks AFB for review and revision. Because of constraint on space, excerpts and condensations from the proposal are provided below to indicate that effort was made to reach the research objective.
  1. Purpose. The intent of the work is to study the effects of pyridostigmine on pilot performance in the AFHRL/OT simulation facility so that usefulness of the drug as an agent of chemical defense for aircrews against toxicants of the organophosphorus anti-cholinesterase series might be further assessed.
  2. Drug. Under the double-blind procedure, the subjects will be treated alike in drug and placebo conditions in a procedure in which a capsule will be given orally,

starting the day prior to a test session. Dosages will be 30 mg of pyridostigmine at intervals of 3 hours, per USAF usage. Tests will be conducted at times and at intervals so that the drug is maximally active in the system at the time of testing and so that there are no carryover effects of the drug from test to test.

3. Subjects. The subjects will be 24 volunteers who are mission-ready pilots and who have at least 500 hours in fighter aircraft and at least 100 hours in their current fighter. The volunteers will be screened at the home base for hyperreactivity to pyridostigmine and again at the Laboratory for potential adverse side effects. The range of the potential side-effects has been identified, and corrective medical measures will be implemented as needed. Further, monitoring, both behavioral and medical, will be in place to correct such adverse reactions as might (rarely) occur.
4. Apparatus. The apparatus of central importance is the flight simulation facility. The flight simulator in this facility will be programmed for four separate tasks which will require 30 minutes for their execution. A nonexhaustive list of behavioral measures in one of the four task components (Red Flag component in which the pilot flies a simulated combat mission) includes the following: number of kills, number of threats detected versus number of reactions to threats, number of rounds fired versus number of threats.

A number of physiological recording devices will be introduced into the flight simulation facility for the purposes of the present work. These measures of stress/arousal reactions in the peripheral nervous system will include the following: integrated body temperature recording, the EKG, pulse plethysmography, respiratory cycle measurement, and the electromyogram.

5. Procedure: At the outset of the study, each subject will separately be given a 30-minute pretest daily for two days. Then, each subject will be entered into a two-factor experiment with repeated measures. In the experiment, all subjects will be tested under conditions in which they individually operate the flight simulator

while under, or not under, the influence of pyridostigmine and while wearing, or not wearing, protective gear. The several behavioral and physiological measures will be taken throughout the subject's performance on the four tasks during the 30-minute session. The subjects will receive a total of four tests, with one separated from the next by a two-day period.

6. Medical monitoring. Informed consent will be obtained by a physician. The informed subject will be given medical and behavioral checks at fixed intervals while he is under the influence of pyridostigmine and at the same intervals after he has ingested the placebo used for control of expectancy effects in the double-blind procedure. Paramedical and medical personnel will be closely involved in the work, medical intervention as needed will be immediately available.
7. Statistical design. A treatments-by-treatments-by-subjects procedure will be used. This procedure tests each subject under each treatment condition within both of the planned factors. This design allows for fewer subjects than is required in factorial design for the same degree of significance. Potential carryover effects must be considered. Therefore, factors are counter-balanced, as are treatments within factors. Also, pyridostigmine will be allowed to clear the body between one test and the next for each subject.
8. Expected outcomes. The major outcome will be integrated medical, performance, physiological, and flight simulator data that can be used by field commanders to evaluate the impact of pretreatment drug in defense against nerve agents.

VIII. COMPARISON OF TELEMETERING VERSUS HARD-WIRED APPARATUS FOR RECORDING PHYSIOLOGICAL EVENTS IN HUMANS:

- a. Approach. Texts, monographs, and journals in two science libraries were examined for discussion of telemetering and biotelemetering instrumentation. On the basis of the literature review, a report was prepared by my research colleague and myself and was submitted to the person in the Laboratory who had asked for the report.

- b. Results. The report is herewith summarized in sentence outline as follows: Telemetry and biotelemetry have come to refer to measurement at a distance of events (biological events) in an organism. Biotelemetric devices may be categorized in several ways - clinical versus research devices, implantable versus attachable, human versus infrahuman monitors. Greatest focus has been on development of attachable ECG devices for monitoring human cardiac functioning. Even where benefits are significant, there are health hazards from RF radiation from the transmitting antenna which must be seen as cautionary to researchers. Also, there are difficulties in recording produced by variables such as the interfering 'noise' of nylon clothing. The latter problem is correctable, but former may not be easily dealt with. Body mounted transmitter antenna radiate omnidirectionally, and the body acts as a reflector and absorber for frequencies above 100 MHz which are well within the range (10 to 1,000 MHz) of radiated frequencies in commercial telemetering apparatus. U.S. standards for this range is  $10\text{mW}/\text{cm}^2$  averaged over 0.1 hour periods. In the USSR, the exposure limit in the same range is lower. Below 300MHz, the upper limit is  $0.01\text{ mW}/\text{cm}^2$  and is averaged over 6 hours/day. Research shows that tissue damage results from 40 microwatts/ $\text{cm}^2$  and 0.1 mWh. Behavioral and physiological dysfunctions may result from power densities as low as a few microwatts/ $\text{cm}^2$ . These dysfunctions include muscular weakness, headaches, visceral problems. Similar problems occur in hyperreactive subjects given pyridostigmine. Telemetry may be a source of confounding in the pyridostigmine study, and its use seems to be contraindicated, even if the measure next interference due to static electricity and the like can be controlled.

#### IX. RECOMMENDATIONS:

- a. Guidance for implementing results of research. Arrangements for funding the study of pyridostigmine effects on pilot performance were made long prior to my minor and transient involvement in the work. What contributions I made were directed toward presenting an experimental design for the study and devising specific procedures for implementation of the investigation. Within the framework of these activities,

I concentrated on choosing appropriate measures of emotional arousal and recommending means for monitoring the physiological correlates of this arousal. The recommendations have been discussed with Laboratory personnel and with, at the direction of the Laboratory members, persons who have contractual arrangements with the Laboratory. An ongoing relationship with both allied groups has been arranged. The sum of \$80,000 has been allocated for development of the monitoring apparatus, and arrangements are being sought with my University for transfer of the funds so that I can continue my association with the project through purchase, testing, and installation of the monitoring equipment.

- b. Suggestion for follow-on research. The USAF is considering feasibility of administering pyridostigmine to aircrews on a sustained basis in times when there is threat of chemical warfare attack. The possible adverse effects of pyridostigmine in multiple treatments, however, is not understood where normal adult subjects are involved. There is evidence that pyridostigmine crosses the blood-brain barrier and that at higher dosages has deleterious effects on behavior. I am planning to submit a Research initiation grant proposal which deals with means to study long-term effects of pyridostigmine on several species of laboratory rodents. The aim of the work is to generate data which may be extrapolated from the experiments with lower mammalian species to humans. Numerous EPA and FDA regulations indicate that such extrapolations have predictive relevance to man.

#### REFERENCES

1. Anon., AMA Drug Evaluation. Philadelphia, Pennsylvania, Saunders, 1975.
2. Anon., Facts and Comparisons, Indianapolis, Indiana, E.J. Lilly and Company, 1983.
3. Bartus, R.T., R.L. Dean, B. Beer, and A.S. Lipps, "The Cholinergic Hypothesis of Geriatric Memory Dysfunction," Science, Vol. 217, pp. 408-417, 1982.
4. Caceres, Cesar A. (Ed.), Biomedical Telemetry, New York, Academic Press, 1965.
5. Calabrese, Edward J., Principles of Animal Extrapolation, New York, New York, John Wiley & Sons, 1983.
6. Carlson, Neil R., Physiology of Behavior, 2nd ed., Boston, Massachusetts, Allyn and Bacon, 1981.
7. Cleary, Alan, Instrumentation for Psychology, New York, New York, John Wiley & Sons, 1977.
8. Goodman, Louis S. and Alfred Gilman (Eds.), The Pharmacological Basis of Therapeutics, 4th ed., New York, New York, The Macmillan Company, 1970.
9. Gosselin, Robert E., Harold C. Hodge, Roger P. Smith and Marion N. Gleason, Clinical Toxicology of Commercial Products, 4th ed., Baltimore, Maryland, The Williams & Wilkins Co., 1976.
10. Harvey, John A., Behavioral Analysis of Drug Action, Glenview, Illinois, Scott, Foresman and Company, 1971.
11. Irwin, S., "Drug Screening and Evaluative Procedures," Science, Vol. 136, pp. 123-128, 1962.
12. Mitchell, Clifford L., Nervous System Toxicology, New York, New York, Raven Press, 1982.
13. Silver, Ann, The Biology of Cholinesterases, New York, New York, Elsevier Scientific Publishing Company, 1974.
14. Venables, Peter H. and Irene Martin, A Manual of Psychophysiological Methods, New York, New York, John Wiley & Sons, Inc., 1967.
15. Wade, Ainley (Ed.), Martindale: The Extra Pharmacopoeia, 27th ed., London, England, The Pharmaceutical Press, 1983.
16. Weiss, Bernard and Victor G. Laties, (Eds.), Behavioral Toxicology, New York, New York, Plenum Press, 1977.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PROPULSION FACILITY PLANNING FOR TEST INFORMATION PRODUCTIVITY  
IMPROVEMENT WITH EMPHASIS ON DATA MEASUREMENT UNCERTAINTY IN THE  
ENGINE TEST FACILITY

Prepared by:	Doyle E. Hasty
Academic Rank:	Assistant Professor of Engineering and Physics
Department and University:	Engineering and Physics Department Motlow State College
Research Location:	Arnold Air Force Station, Aeropropulsion Programs Department, Technology Division, Propulsion Analysis Group
USAF Research:	Dr. Grant T. Patterson
Date:	August 3, 1984
Contract No:	F49620-82-C-0035



PROPULSION FACILITY PLANNING FOR TEST INFORMATION PRODUCTIVITY  
IMPROVEMENT WITH EMPHASIS ON DATA MEASUREMENT UNCERTAINTY IN THE  
ENGINE TEST FACILITY

by

Doyle E. Hasty

ABSTRACT

High subsonic, transonic, and supersonic propulsion system testing at one of the most unique aerospace testing facilities in the world, the Engine Test Facility (ETF), requires that maximum use be made of today's increased computational power, advanced information management systems and high-speed digital data acquisition systems. The heart of ETF was created in Germany in 1944 as the Bavarian Motor Works (BMW) which was later dismantled after the war and shipped to the United States. The early 1950's supported construction of ETF.

Various expansions and adaptations of this facility have continued for over thirty years to the point that ETF has not only played a key role in the development of many of the United States' aerospace programs, but will continue to play that role for decades to come if state-of-the-art modernization can be designed and implemented into this facility.

The objective of this research analysis is to evaluate new test techniques, data acquisition processes, data reduction techniques, data displays and analysis techniques to improve and modernize the propulsion test information process. A special emphasis was to analyze methods for improving and monitoring the data accuracy or data uncertainty of various test parameters.

### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very educational summer at the Arnold Engineering Development Center, Tullahoma, Tennessee. He would like to thank Sverdrup Technology, Inc., for their hospitality, excellent working environment, and challenging assignment.

Additionally, he would like to thank Dr. Grant T. Patterson for suggesting this area of research and for his collaboration and assistance. He would like to acknowledge the assistance and cooperation of Dr. Virgil K. Smith III, Mr. Marshall Kingery, Mr. Alvis Turrentine, and Mr. Shirley Williams who gave of their time to make this such a meaningful assignment.

## I. INTRODUCTION:

The updating and modernization of high-altitude turbojet engine and rocket test cells presently concerns the Air Force Systems Command at the Arnold AF Station, Tennessee, because of the aging facilities that exist. As these test chambers begin their fourth decade of testing service, state-of-the-art computing power, information systems, and revised testing approaches must be incorporated into the major existing test chambers.<sup>1</sup> The author has had 15 years of experience in high-altitude aerospace testing and found this particular assignment very challenging and well matched to his experience and abilities.

The broad range of experience gained by research and analysis of a project of this scope was extremely beneficial to the author. He was able to quickly become productive in this assignment and accomplish more than is usually normal with a project that requires a given adaptation period. As head of the engineering department at his school, the author will use this state-of-the-art technology experience in designing and selecting new engineering and physics laboratory and analysis equipment for the school's new engineering building which is presently under construction.

The author feels that the opportunities afforded by this research effort at this time will be immeasurably beneficial to his engineering students, many of whom eventually will become engineers permanently employed at this Air Force base or may even work here cooperatively as they finish their engineering degrees.

## II. OBJECTIVES:

The goals selected for this research effort included the analysis and evaluation of presently used testing techniques in the high-altitude turbojet engine and rocket test cells. The major goal was to evaluate state-of-the-art test techniques, data acquisition processes, data reduction techniques, data displays, and analysis techniques in order to improve propulsion test information productivity. The final goal was to recommend approaches for improving productivity by considering hardware change, new or common software for data reduction/communication and changes in testing philosophy to accommodate new hardware and software. Finally, estimates of the data uncertainties from any modernized system must be determined.

## III. PRESENT TESTING TECHNIQUES IN HIGH-ALTITUDE TEST CELLS:

Turbojet engine and rocket test cells require that the test articles and associated test cell systems be controlled safely and precisely to a set of pre-planned test conditions. These environmental and test article conditions, once produced, must be maintained while data acquisition is in progress. A centralized area consisting of the test cell, surrounding test area, and control room are dedicated to ensuring that these conditions are met and maintained both reliably and safely.

During testing, the control room personnel and equipment are used primarily to control the test article and associated test cell equipment to the required test mission conditions. Presently, the environmental conditions, however, are set and maintained at another location by personnel and equipment dedicated to that task. Interaction and coordination between these two areas are primarily audio and alphanumeric display communications, indicating the environmental set points and conditions.

Interactions with the test article and its on-board system are performed primarily manually. The control and interaction are maintained by observing visual indications and performing manual responses.

#### IV. PROPOSED IMPROVEMENTS IN TESTING TECHNIQUES:

Analysis of existing testing conditions led to the conclusion that the following four major improvement areas are of the highest priority.<sup>2</sup>

- A. Reduce the testing time required for each test article.
- B. Reduce the testing cost of each test article.
- C. Increase the availability of each test cell.
- D. Meet needs of future complex computer controlled test articles, i.e., spacecraft, spaceships, and space stations.

This research effort provided the basic engineering study and analysis to produce a preliminary design plan. The plan that was developed included using state-of-the-art techniques to minimize the time in obtaining test article and test cell conditions necessary to acquire data. Possible automation of the following control system functions were determined to be necessary when analyzing techniques to minimize this time frame.

- A. Throttle control
- B. Sequence control
- C. Health monitoring control
- D. Operation interface with engine control
- E. Emergency control
- F. Event logging control
- G. Auxiliary system controls (test cell cooling air, exhaust duct cooling, lab seal control, and fuel flowmeter selection)
- H. Systems readiness between environment controls and data acquisition computer.

It is now anticipated that three levels of automation complexity will exist. These are total automation, partial automation, and no automation. Those systems evaluated for total automation, hands-off operation, will be systems that can be set up prior to test time and will operate under computer control during the test. Partially automated systems are those

where limited manual inputs will be necessary during testing. Systems classified as no-automation will function entirely manually. The primary consideration for manual systems will involve safety of personnel and equipment.

In determining the level of automation required, the following human factors will be investigated.<sup>3</sup>

- A. Human/machine interface.
- B. Dedicated/distributed system.
- C. Safety/reliability.
- D. Efficiency/time savings.

The hardware needed to carry out each controls, data acquisition, and data processing function is being identified. This hardware will include state-of-the-art equipment as actuators, controllers, convectors, transducers, data acquisition systems, and data processors. After the computer hardware is identified, the associated software requirements will be specified.

The data uncertainties of the data system including the data acquisition and data processing must be assessed as part of the detailed design effort. Presently, the Engine Test Facility uses two standards as guides to establish data measurement uncertainties.<sup>4,5</sup> Individual system models of some of the proposed systems with predicted data uncertainties were performed by the author as part of this research effort.<sup>6</sup>

The modernization and updating is proposed for the six major test cells in the Engine Test Facility. The preliminary design estimates show an anticipated savings for each test unit in excess of \$750,000. The preliminary cost estimate of modernization per test unit is in the \$1,000,000 range ensuring a less than two-year payback period.

This research period covered the preliminary design phase which will precede the now-to-be-completed detailed design phase, implementation, and checkout in a selected model test cell over the next year.

V. RECOMMENDATIONS:

The author proposes to continue this effort through a Research Initiation grant so that he can be involved with the detailed design phase, implementation, and checkout of a modernized model test cell during the next year. This year-long activity would complete the design effort and complete the implementation in the first test cell. Then this design plan could be used for the five other major test units to give the Engine Test Facility a state-of-the-art testing technique with modernized computer-controlled facilities and equipment.

The author feels that this summer research effort has been very beneficial to him as well as to the future of high-altitude turbojet engine and rocket testing at the Arnold Engineering Development Center. He looks forward to being able to see this research effort totally completed and implemented to prepare the Engine Test Facility for testing into the twenty-first century.

#### REFERENCES

1. Test Facilities Handbook, Arnold Engineering Development Center, Air Force Systems Command, United States Air Force, March 1984.
2. Welch, Jack L. and Jerry L. Brinkley, "Automating the Test Area Controls Systems of the Turbojet Test Units of the Engine Test Facility," AEDC No. EA44, CECORS No. 820072-1, September 1983.
3. Rosinger, George, "Preliminary Human Factors Standards, Specification, and Guidelines for Engine Testing Control Room," Battelle Laboratories, Columbus, Ohio, September 1983.
4. Abernethy, R. B. and J. W. Thompson, Jr., "Handbook, Uncertainties in GAS TURBINE MEASUREMENTS," AEDC-TR-73-5 (AD755356), February 1973.
5. Abernethy, R. B., "Interagency Chemical Rocket Propulsion Group Handbook for Estimating the Uncertainty in Measurements Made With Liquid Propellant Rocket Engine Systems," CPIA Publication No. 180, April 1969.
6. Hasty, D. E. and S. C. Williams, "Steady-State Measurement Uncertainties," Engine Test Facility, Arnold Engineering Development Center, July 1984.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A STUDY ON POINT DIFFRACTION INTERFEROMETRY

PLUS

HOLOGRAPHIC MEASUREMENTS OF A TURBULENT BOUNDARY LAYER

ON A ROUGHENED WIND TUNNEL WALL

Prepared by:	Dr. A. George Havener
Academic Rank:	Assistant Professor
Department and	Department of Mechanical Engineering
University:	University of Dayton, Dayton, Ohio
Research Location:	Air Force Wright Aeronautical Laboratories, Air Force Flight Dynamics Laboratory, Aeromechanics Division, Instrumentation Branch, Aero-Optic Instrumentation Group
USAF Research:	Dr. George L. Seibert
Contract No:	F49620-82-C-0035

A STUDY ON POINT DIFFRACTION INTERFEROMETRY  
PLUS  
HOLOGRAPHIC MEASUREMENTS OF A TURBULENT BOUNDARY  
LAYER ON A ROUGHENED WIND TUNNEL WALL

by

A. George Havener

ABSTRACT

The general theory of a point diffraction interferometry is discussed along with a description of a specific system for direct application to wind tunnel testing. Problems involved with the use of a point diffraction interferometer, and plans for continued developments and applications plus recommendations for construction of an improved system are also presented. Additionally, holographic interferometric measurements of a two dimensional, turbulent boundary layer on a roughened wall of a Mach 3 high Reynolds wind tunnel are presented in direct comparison to similar measurements of a turbulent boundary layer on a smooth wall of the same facility. Plans for detailed reduction of these interferometric data to density and velocity distributions are also discussed.

#### ACKNOWLEDGEMENT

The author extends thanks and appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him the opportunity to accomplish the research presented herein, and for the opportunity to work and study in the professional environment of the Instrumentation Branch of the Air Force Flight Dynamics Laboratory at Wright-Patterson Air Force Base, Dayton, Ohio. Special thanks and gratitude are extended to Dr. George L. Seibert, Mr. Clifford B. Weissman and Mr. C. Dean Miller of the Aero-Optic Instrumentation Group for their assistance throughout this project, and for their continued friendship over many years of professional association.

## I. INTRODUCTION:

For the past fifteen years, the author has been involved with the design and application of optical measuring and flow visualization techniques to the study of compressible fluid flow fields. The dominant technique has centered on holography with the emphasis being holographic interferometry (Ref 1). Early applications resulted in acquiring unique density data of two dimensional, supersonic turbulent boundary layers for both attached and separated states, and recent research has been completed wherein holographic interferometry was used to obtain density data of turbulent bursts in an axisymmetric laminar boundary layer on a sharp tip cone in a Mach 6 flow field; both laminar and turbulent distributions of density were also obtained (Ref 2). Additionally, the author has developed a computer approach to studying the interferometric depiction of theoretical flow fields. This process, called computational interferometry, allows for the detailed interferometric visualization of theoretical flow fields obtained from numerical solutions to the Navier Stokes equations; the computer generated interferograms are the results that would be seen if the flow field were actually viewed using a real interferometer.

The nature of the present USAF research addresses the problem of obtaining real-time density data of wind tunnel

tests. While many successful interferometric applications to wind tunnel studies have been obtained during the past decade, some of the major problems in the general and routine use of interferometry have not been solved. Among these problems, the three most prevalent are data reduction (automated fringe reading capabilities are needed), multi-viewing (multiple views of flow fields are needed to permit resolve of three dimensional flows), and rapid recording of interferograms (typically, recording rates of 10 to 20 thousand interferograms per second are needed to obtain data of transient phenomena). Recent developments of a new interferometric technique, Point Diffraction Interferometry (PDI:Ref 3-5) and its application to aero-thermodynamic studies (Ref 6) show definite promise for obtaining real-time interferometric data of wind tunnel tests. Coupled with a reliable automated fringe reading device, one can imagine an optical instrument capable of obtaining instantaneous density data of two dimensional flows, and if the problem of multi-viewing were also solved, one can further recognize the ability to quantify the density fields of almost any flow that is translucent to light waves.

## II. OBJECTIVES OF THE RESEARCH EFFORT:

The primary objective of this research effort is to verify the usefulness and the accuracy of Point Diffraction

Interferometry applicable to wind tunnel testing. A secondary objective of this research is to obtain interferometric measurements of an on-going wind tunnel test on surface roughness effects in turbulent, supersonic boundary layers. Noteworthy here is the fact that Bachalo and Houser (Ref 6) presented proof on June 25-27, 1984 that PDI can be used successfully in wind tunnel applications, and therefore, the primary objective of this research shifted from an exploratory effort to one of development and application of an In-House PDI system at the Air Force Flight Dynamics Laboratory. In this sense, the effort has been directed towards obtaining practical knowledge and experience on the use of PDI by building a working system from available Air Force optical equipment, and subsequently applying this system to actual aero-thermodynamic tests.

### III. CONCEPTS AND GENERAL THEORY OF PDI:

Fundamentally, interferometry involves the blending of coherent sets of light waves in a controlled way so that a meaningful interference (the interferogram) is constructed. As used here, meaningful interference is taken to mean an interference from which measurable distortions to one of the sets of light waves can be determined with certainty. Classical interferometric applications to wind tunnel test are not generally done, because satisfying the strict

optical requirements of these devices is too difficult. Laser applications relax these requirements considerably so that interferometry--specifically holographic interferometry--can be done with relative ease at most wind tunnel facilities. However, these optical processes require the use of an optical reference (this is generally referred to as the reference beam) which makes acquiring real-time data very difficult; indeed, in most facilities, acquiring real-time interferometric data is practically impossible. Different from these interferometers, PDI offers to opportunity to acquire real time interferometric data, because the optical reference is not a separate beam of light that traverses long optical paths around the wind tunnel facility. Instead, the optical reference of the PDI is generated from the test beam (the test beam is the beam of light waves that pass through the test section of the wind tunnel), and this generation of reference waves can be done continuously so long as the test beam exists.

The fundamental concept of the PDI involves generating an optical reference from a set of disturbed light waves, and subsequently using this reference to produce meaningful interference with the test waves. References 5 and 6 show ways in which PDI can be used with a single pass Toepler schlieren system. As shown in Figure 1, the schlieren filter is replaced by a collimating lens and a small Mach-

Zehnder interferometer. In use, the beam of light traversing the test section of the wind tunnel is captured and imaged by the second schlieren head, recollimated by the positive lens and directed into the small Mach-Zehnder interferometer. Once in this interferometer, one portion of the separated beam is spatially filtered by a lens-pin hole assembly while the other portion of the beam is held constant. At the second splitter plate of the interferometer, the filtered and unfiltered beams are reunited to form an interference.

The spatial filter in the Mach-Zehnder inteferometer establishes the optical reference for the interference. Light waves traversing a compressible flow in a wind tunnel are refracted by the density field of the flow. When these light waves are focussed by a positive lens, the refracted parts of the waves are focussed to different spots in the focal plane of the positive lens while all the nonrefracted portions of the waves are focused to a central spot. By carefully inserting a pin hole in the focal plane of this lens, all the refracted portions of the light waves are blocked--if the pin hole is small enough, diffraction of these light waves will even remove optical information of the silhouettes of the solid boundaries of the wind tunnel so that the waves exiting the filter are seen as pure spherical waves diverging from a point source--and the



light passing through the filter becomes the optical reference for the interferometer. Adjustment of the resulting interference is regulated by traversing the pin hole relative to the focussing lens. The second lens on the "down stream" side of the pin hole is used to recollimate the diverging light waves. The lens pair of this spatial filter invert the wave fronts in this path of the interferometer which can produce a visible and somewhat annoying effect seen as a double image of the test subject on the viewing screen. To negate this double image effect, a lens pair similar to that of the spatial filter is placed in the test path of the interferometer.

Both continuous and pulsed light sources can be used with PDI. One recognizes that by using a pulsed light source, the improvement of PDI over the use of holographic interferometric techniques is not significant. However, the use of a continuous light source presents a breakthrough in applied interferometry, because it now appears possible to obtain continuous, real-time interferometric measurements of wind tunnel flow fields. Conceptually, this is done by imaging the light waves exiting the interferometer into the focus of a high speed camera.

At this time, the advantages of PDI appear to be:

1. Acquisition of real-time interferometric data.
2. Use of relatively simple optical system.

3. General application to wind tunnel facilities  
with existing schlieren systems.

IV. FABRICATION AND TESTING OF AN IN-HOUSE PDI:

A small PDI has been built and preliminary tests have been performed in the laboratory of the Aero-Optic Instrumentation Group of the Air Force Flight Dynamics Laboratory. This system is powered by a 50 mw continuous wave krypton laser. The schlieren path is established by two on-center parabolic mirrors (dia. of 229 mm,  $f/5$ , nominally), and a 10X microscope objective is used to diverge the beam from the source. The schlieren path is off axis by approximately 7.5 degrees, and consequently, an aberation to the wave fronts is produced. This aberation is negated by placing a cylindrical lens in the focus of the microscope objective; the aberation has to be negated before the light waves pass through the spatial filter in the interferometer so that proper spatial filtering is accomplished. The Mach-Zehnder interferometer consists of two 3 degree prism beam splitters and two first surface plane mirrors. These components are mounted to standard triangular optical rails, and the entire system is secured to an air cushioned vibration damped optical table. In the normal use of a PDI at a wind tunnel facility, the schlieren heads, light source and recording optics would be

mounted on separate stands, and therefore, this small system is built to simulate these conditions; the schlieren mirrors, interferometer and recording optics are mounted to separate stands. The spatial filter of this PDI consists of a 50 micron pin hole and a pair of  $f/3$  lenses (f.l. of 150 mm, nominally), and the beam diameter of the light waves in the interferometer is approximately 25 mm.

Thusfar, the results obtained are marginally successful. Procedures for properly aligning the PDI--considerable care is needed to correctly align the spatial filter--have been learned from numerous attempts to "get it right," and still only poor quality interferograms have been obtained. All of the problems seem to center on not having the proper equipment. Specifically, the optical components of the interferometer and those of the spatial filter assembly are adequate for demonstrating that a PDI system of this type works, but they seem to be inadequate for obtaining and maintaining good alignment. The lens-pin hole combination of the spatial filter is not correct (the present one is the best that could be put together with existing equipment), nor is the match between the cylindrical lens and the off-axis angle of the schlieren path correct. Consequently, the effect of the aberration of the light waves is focussed by the spatial filter, and further distorted by diffraction at the pin hole. This results in the formation

of nonuniform, asymmetric fringes, and this focussed aberation inhibits generation of good optical reference waves.

Two other noteworthy problems encountered thusfar are light intensity and system stability. The spatial filter appears to block approximately 80 percent of the incident radiation which means the intensity of the light coming from the filter is very low. Good fringe contrast results when the intensity of both light beams is nearly equal, and therefore, the intensity of the unfiltered waves in the interferometer has to be reduced. For the present system, this is generally accomplished by using a 90-10 beam splitter at the entrance to the interferometer, and using the spatial filter in the 90 percent path. However, for the present system, a relatively low power laser is being used, and fringe distinction is very difficult particularly in the image plane of the interferometer. This problem appears to be easily resolvable by using a more powerful laser (e.g., a 2 or 3 watt cw laser).

System stability presents concern for the general use of PDI especially in "noisy" wind tunnel complexes. As previously stated, the major components of this small PDI are mounted on separate stands so as to simulate normal conditions of an actual wind tunnel facility. The optical paths of this system are short and unprotected; no special

shielding is used and the light waves travel through room air. Apparently the single most dominant factor affecting the formation of the fringes in a PDI system is the relative alignment of the spatial filter. Any vibration altering this alignment from the true setting causes a distortion in the generated reference light waves which can ruin the interference. Since the nominal size of the pin hole is on the order of 50 microns, maintaining continuous alignment of the spatial filter places strict stability requirements on the design of the PDI system. For the present system, both the overhead room air conditioner or a sharp foot stomp on the floor were found to be strong enough disturbances to offset the alignment of the spatial filter.

At the present, work is continuing on the development and use of the in-house PDI system. Improved optical components are being sought, and a convective heat transfer model is being developed to serve as an initial physical application. Using this model, PDI interferograms will be obtained and analyzed for measuring accuracy and sensitivity. Following this, the PDI system will be applied to a wind tunnel study to further verify its usefulness.

#### V. APPLICATION OF HOLOGRAPHY TO WIND TUNNEL TEST

Pulsed laser holography is used to support a study the

turbulent boundary layer on both a smooth wall and roughened wall of a Mach 3 high Reynolds wind tunnel. The holographic laser system used in this application is the Aero-Optic Instrumentation's pulsed laser system which is well documented in the literature (Ref 1), and the details of this system are not discussed here.

The objective of this effort is to obtain detailed density data of these boundary layers by reducing holographic interferometric measurements of the flows. Both double exposure--the same hologram is exposed twice, once with flow and once with no flow in the wind tunnel--and dual plate holography--two separate holograms are used to record the flow and no flow conditions--are used to obtain the interferograms. Both of these techniques are also discussed in the literature (Ref 1).

Typical results of this research are shown in Figures 2 and 3. Figure 2 is direct comparison of the smooth and roughened wall cases for the same stagnation pressure, and as shown, the boundary layer for the roughened wall is thicker and appears to contain more local turbulence than the smooth wall flow. Furthermore, the free stream flow fields seen in these two interferograms are not the same; again, the roughened wall case shows the presence of more turbulence. The bumps along the roughened wall are known to generate turbulence, but as shown in the interferograms,

this generated turbulence appears to be transmitted to the free stream where it undoubtedly cause an additional disturbance to the boundary layer. The possibility of this turbulence being reflected from the other walls of the wind tunnel is also real. Figure 3 shows a comparison between interferograms for the roughened wall boundary layers for two different stagnation pressures. As seen, the fringe definition for the higher pressure shows a much increased phase shift across the boundary layer, and this suggests a limit on the use of holographic interferometry to study these boundary layers at high stagnation pressures. Thusfar, no limit has been found in the present study. Also shown in Fig 3 is a double pulse holographic interferogram. This interferogram is made by exposing the flow field twice; the two exposures are temporally spaced by approximately 200 micro-seconds which means the fringes of this interferogram show the existence of unsteady behavior of a frequency exceeding 5 KHz.

The optical results of this effort have been obtained only recently, and time does not permit the detailed quantitative data available from analysis of the holograms to be included in the report. These data are being obtained at this time, and will be presented elsewhere (e.g., AIAA paper and/or AFWDL AFFDL/TR) in the future.

## VI. RECOMMENDATIONS:

Based on the research accomplished to date, the following recommendations are made:

1. An improved PDI system needs to be constructed using optical components specifically designed for use in the PDI system. More testing and evaluation of the test results need to be done so that the concerns for stability and illumination power that were found in this effort can be analyzed properly. While inconclusive at this time, PDI still appears to be a worthwhile process for obtaining continuous and real-time interferometric measurements of wind tunnel tests.
2. The assembled in-house PDI system should continue to be used so that improved understanding and practical experience in PDI applications and measuring techniques can be obtained.
3. Reduction of the holographic data of the roughened wall turbulent boundary layer study should be done as quickly as possible. These data are new and unique, and offer the possibility to obtain meaningful quantitative information throughout the boundary layer.



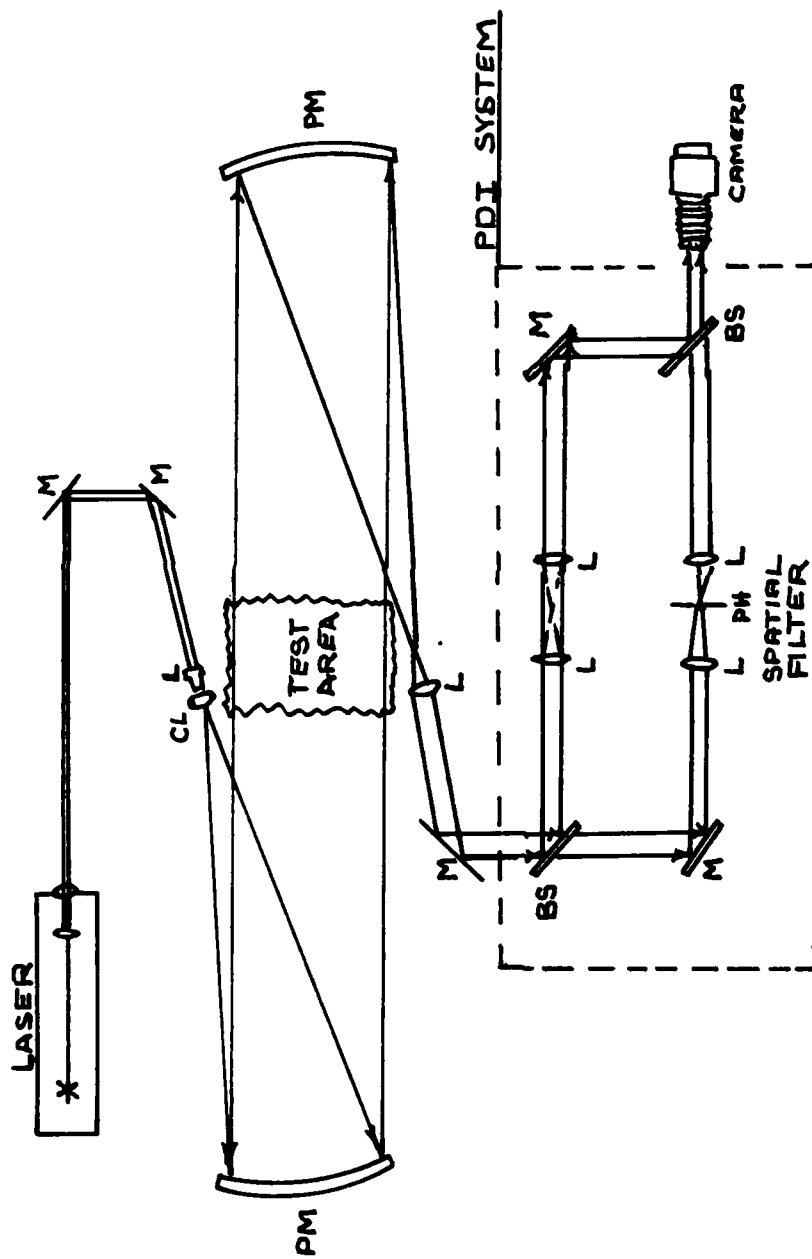
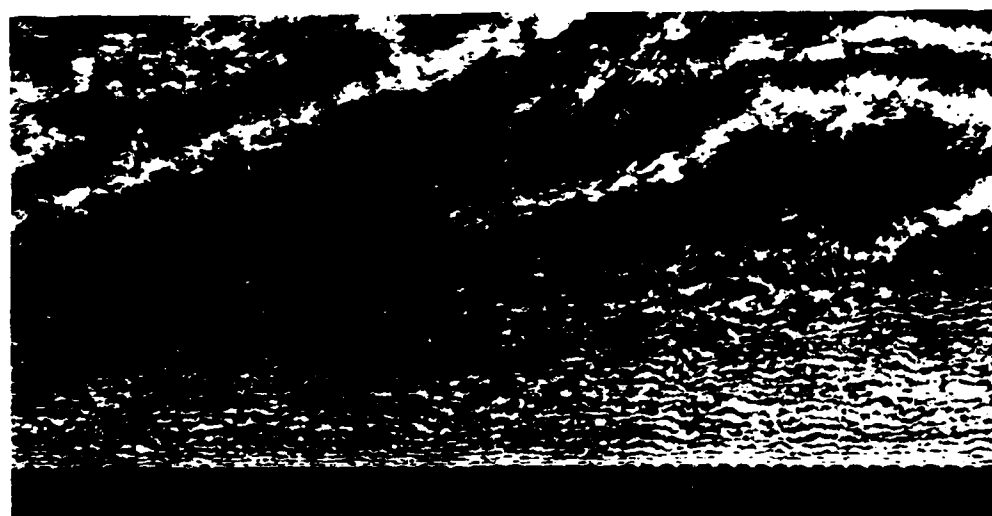


FIG 1 SCHEMATIC DIAGRAM OF A POINT DIFFRACTION INTERFEROMETER FOR WIND TUNNEL APPLICATIONS



SMOOTH WALL (FRINGE SHIFT = 16)



ROUGH WALL (FRINGE SHIFT 22)

FIG 2 INTERFEROGRAMS OF 2-D TURBULENT BOUNDARY LAYERS ON SMOOTH AND ROUGHENED WALLS OF A MACH 3 WIND TUNNEL ( $P_0 = 586$  KPa.)



STAGNATION PRESSURE 586 KPA.  
(FRINGE SHIFT = 22)



STAGNATION PRESSURE 999 KPA.  
(FRINGE SHIFT = 37)



DOUBLE EXPOSURE INTERFEROGRAM  
( $\Delta T = 200 \mu s$ )

FIG 3 INTERFEROGRAMS OF A 2-D TURBULENT BOUNDARY LAYER ON A ROUGHENED SURFACE

## REFERENCES

### Conference and Journal Publications:

1. Havener, A.G., "Users Guide on Pulse Laser Holography for Wind Tunnel Testing," ARLTR 75-0213, Aerospace Research Laboratories, AD/A 017710, June, 1975.
2. Havener, A.G., "Holographic Measurements of Transition and Turbulent Bursting in Supersonic Axisymmetric Boundary Layers," AIAA Paper No 83-1724, AIAA 16th Fluid and Plasma Dynamics Conference, Danvers, MA, July, 1983.
3. Smartt, R.N., "Special Applications of the Point-Diffraction Interferometry," SPIE, Vol. 192, Interferometry, pp 35-40, 1979.
4. Howes, W.L., "Large-Aperture Interferometer With Local Reference Beam," Journal of Applied Optics, Vol. 23, No. 10, pp 1467-1471, May, 1984.
5. Howes, W.L., "Large-Aperture Interferometer Using Local Reference Beam," NASA-TP-2060, 1982.
6. Bachalo, W.D. and Houser, M.J., "A Real-Time Interferometer Technique For Compressible Flow Research," AIAA Paper No. 84-1600, AIAA 17th Fluid Dynamics, Plasma Dynamics, and LAser Conference, Snowmass, Colorado, June, 1984.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EFFECTS OF TEMPERATURE AND REACTANT SOLVATION UPON  
THE RATES OF GAS-PHASE ION-MOLECULE REACTIONS

Prepared by:	Peter M. Hierl
Academic Rank:	Professor
Department and University:	Department of Chemistry University of Kansas
Research Location:	Air Force Geophysics Laboratory (LID) Hanscom Air Force Base, Massachusetts
USAF Research:	Dr. John F. Paulson
Date:	August 25, 1984
Contract No.:	F49620-82-C-0035

EFFECTS OF TEMPERATURE AND REACTANT SOLVATION UPON  
THE RATES OF GAS-PHASE ION-MOLECULE REACTIONS

by

Peter M. Hierl

ABSTRACT

The rate constants and the product branching ratios for the nucleophilic displacement reactions of  $\text{CH}_3\text{O}^-$  and  $\text{OH}^-(\text{H}_2\text{O})_n$  (where  $n = 0, 1$ , or  $2$ ) with the methyl halides, and for the proton transfer reactions of  $\text{CH}_3\text{O}^-$  and  $\text{OH}^-(\text{H}_2\text{O})_n$  with acetonitrile and the hydrogen halides have been measured in the gas phase over the temperature range 200-500 K, using the AFGL selected ion flow tube (SIFT). The rate constants of the fastest reactions were found to be very close to the collision rate. Reactant solvation was found to decrease the rates of the nucleophilic displacement reactions, most significantly in the case of the least exothermic reactions, but was found to have little effect upon the rates of the proton transfer reactions. Likewise, increased temperatures decreased slightly the rates of the nucleophilic displacement reactions but had less effect upon the rates of the proton transfer reactions.

#### ACKNOWLEDGEMENT

The author would like to thank the Air Force System Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a stimulating and productive summer at the Air Force Geophysics Laboratory (LID), Hanscom AFB, MA. He would like to acknowledge the laboratory for its hospitality and excellent working conditions.

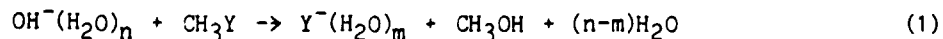
Finally, he would like to thank Dr. John F. Paulson for his valuable suggestions and his guidance, and to thank his colleagues, Prof. Michael J. Henchman, Dr. Albert A. Viggiano, Mr. Fred Dale and Mr. Anton Ahrens, for their invaluable assistance.

## I. INTRODUCTION

Gas-phase reactions between ions and neutral molecules play a significant role in atmospheric chemistry, radiation chemistry, combustion chemistry, and chemical processes occurring in lasers and in electrical discharges. Moreover, studies of ion-molecule reactions can provide important basic data on the energetics and mechanisms of gas-phase reactions in general, as well as provide valuable insight into ionic reactions occurring in solution. Consequently, gas-phase ion chemistry has been an active and rapidly growing field of chemical research for the past several decades.<sup>1</sup>

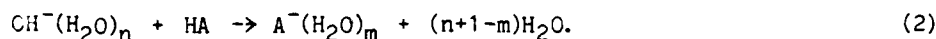
An important recent advance in this field has been the development of experimental techniques by which selectively solvated ions can be formed and the resulting effect of solvation upon the ion's reactivity can be studied.<sup>2</sup> These studies have been motivated both by the recognition that solvated ions play an important role in the chemistry of the lower ionosphere,<sup>3</sup> and by the hope that such studies might help elucidate the nature of the solvent effects which often dominate the rates of ionic reactions in solution.<sup>2,4,5</sup>

The present project is primarily concerned with investigating the effects of solvation and of temperature upon the reactivity of the hydroxide anion, OH<sup>-</sup>, in the gas phase. Two common types of ionic reactions were chosen for this study: (1) nucleophilic displacement (S<sub>N</sub>2) reactions between hydrated OH<sup>-</sup> and the methyl halides, which can be represented by the general equation





where  $n = 0, 1, 2$ , or  $3$ , and  $X = F, Cl, Br$ , or  $I$ ; and (2) proton transfer reactions between hydrated  $OH^-$  and various proton donors, including the hydrogen halides,



The rate constants for the nucleophilic displacement reactions of the solvated hydroxide anion with  $CH_3Cl$ ,  $CH_3Br$ , and  $CH_3I$ , have been measured in a flowing afterglow at room temperature by Bohme and coworkers,<sup>2,5-7</sup> who also measured room temperature rate constants for a number of proton transfer reactions involving  $OH^-(H_2O)_n$ .<sup>2,8,9</sup> We have measured, using the AFGL tandem mass spectrometer, integral cross sections and product branching ratios as a function of reactant translational energy ( $0.03 - 5$  eV center-of-mass) for the reactions of  $OH^-(H_2O)_{0,1,2}$  with  $CH_3Cl$  and  $CH_3Br$ .<sup>10</sup>

## II. OBJECTIVES OF THE RESEARCH EFFORT

The objectives of this project were to investigate the following phenomena for the reactions described in Section I:

- (1) The effect of solvation ( $n = 0, 1, 2$ , or  $3$ ) upon the rate constant of the reaction.
- (2) The effect of temperature upon the rate constant.
- (3) The degree to which the product anion is solvated (i.e., the relative abundances of the various possible products  $Y^-$ ,  $Y^-(H_2O)$ , etc.) in a given reaction.
- (4) The effect of reaction exothermicity (varied by changing the identity of the halogen atom  $Y$  or  $A$ ) upon each of the three previously listed phenomena.

### III. EXPERIMENTAL

The rate constant measurements were carried out using the AFGL Selected Ion Flow Tube (SIFT) apparatus, which is similar to that described by Adams and Smith.<sup>11</sup> Briefly, the reactant ions are generated in a high-pressure ion source, mass-selected by a quadrupole mass filter, and then injected into the flow tube through a venturi-type aspirator. The injected ions are brought into thermal equilibrium through multiple collisions with the carrier gas (He or H<sub>2</sub>), the temperature of which can be varied over the range 200K - 500K. The thermalized reactant ions are then carried downstream past the inlet where the neutral reactant is admitted at a measured flow rate. Further downstream ionic reactants and products are mass analyzed by a second quadrupole mass spectrometer and are detected by conventional pulse counting techniques.

Data for a particular reaction at a given temperature are obtained for a number N (typically, N = 3-10) of replicate runs in each of which the intensity of the reactant ion is measured at eight different flow rates of the neutral reactant. The total rate constant for the selected ion-molecule reaction is determined from the decrease of the reactant ion signal as a function of the neutral reactant's concentration.

Calibration of the neutral reactant flow meter showed that the actual flow rate was approximately 20% less than that indicated by the meter, when the most sensitive scale of the meter was used; no error was found on the other 2 scales. Consequently, rate constants reported here for those experiments in which the flow meter was operated on the most sensitive scale have been multiplied by 1.2 to correct for the error in the flow measurement.

#### IV. RESULTS AND DISCUSSION

##### A. CALIBRATION EXPERIMENTS

To test the accuracy of the techniques employed in this experiment, rate constants were measured at room temperature (298K) for several negative ion-molecule reactions for which reliable rate constant measurements have been reported from other laboratories. In most cases, calibration reactions were chosen which involved the same neutral reactants to be studied in the present work. Table I lists these calibration experiments and compares the present results with those obtained elsewhere.

TABLE 1. Rate Constants for Calibration Reactions at 298K.

Reaction	Rate Constant ( $10^{-10}$ cm <sup>3</sup> /molec-s)	
	This-Work	Other Work
(1) $F^- + HCl \rightarrow HF + Cl^-$	15.6	15.5 <sup>a</sup>
(2) $F^- + HBr \rightarrow HF + Br^-$	11.9	12.4 <sup>a</sup>
(3) $F^- + HI \rightarrow HF + I^-$	7.0	10.1 <sup>a</sup>
(4) $F^- + CH_3Cl \rightarrow CH_3F + Cl^-$	18.1	8.0 <sup>b</sup> 18.0 <sup>c</sup> 19.0 <sup>d</sup>
(5) $F^- + CH_3Br \rightarrow CH_3F + Br^-$	22.2	6.0 <sup>b</sup> 12.0 <sup>d</sup>
(6) $F^- + CH_3I \rightarrow CH_3F + I^-$	23.4	
(7) $OH^-(H_2O) + CO_2 \rightarrow HCO_3^- + H_2O$	8.3	6 <sup>e</sup>
(8) $OH^-(H_2O) + SO_2 \rightarrow HSO_3^- + H_2O$	18.1	20 <sup>e</sup>

<sup>a</sup> Reference 12

<sup>b</sup> Reference 13

<sup>c</sup> Reference 6

<sup>d</sup> Reference 14

<sup>e</sup> Reference 15

In most cases the present results agree very well with the previously reported rate constants; there are, however, several noteworthy exceptions. First, the rate constant measured for Reaction (3) is about 30% less than that reported by Weishaar *et al.*<sup>12</sup> It is possible that this discrepancy is caused by impurities in the HI used in the present study, and analysis of this HI should be made. Second, the rate constant measured for Reaction (4) agrees very well with that obtained by Bohme and coworkers<sup>6,14</sup> from a flowing afterglow experiment, but is much larger than the value obtained by Olmsted and Brauman<sup>13</sup> in an ion-cyclotron resonance (ICR) experiment. This discrepancy raises the possibility that the reactant ions in an ICR cell do not achieve true thermal equilibrium. Third, the rate constant measured for Reaction (5) is considerably larger than the values previously reported by either Bohme<sup>14</sup> or Olmsted and Brauman.<sup>13</sup> Although the discrepancy with the ICR result might be rationalized by the failure to achieve thermal equilibrium in the ICR cell, the discrepancy with Bohme's flowing afterglow measurement is less easy to explain.

#### B. REACTIONS OF $\text{OH}^-(\text{H}_2\text{O})_n$ WITH THE HYDROGEN HALIDES, METHYL HALIDES, AND ACETONITRILE

The major portion of this project was concerned with the temperature dependence of the rate constants and product branching ratios for the reactions of the solvated anion  $\text{OH}^-(\text{H}_2\text{O})_n$  with the hydrogen halides, the methyl halides, and acetonitrile. The reactant systems studied and the number of temperatures at which rate constants were measured for each system are summarized in Table 2.

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

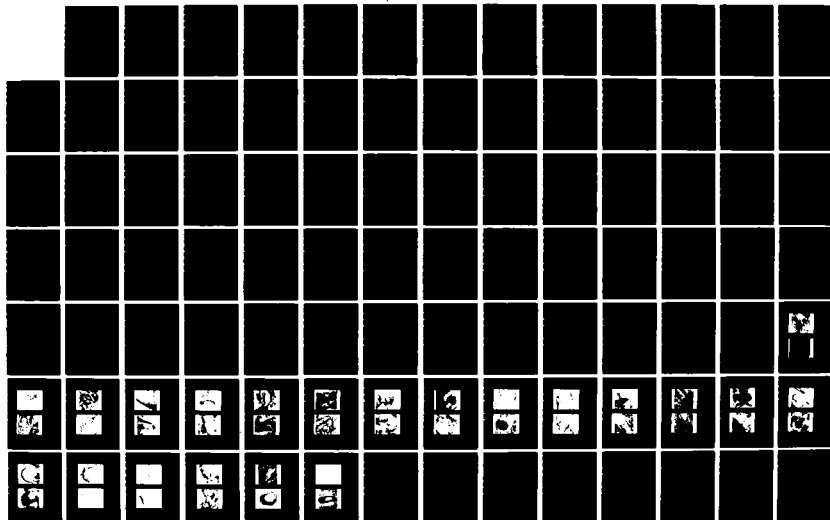
4/13

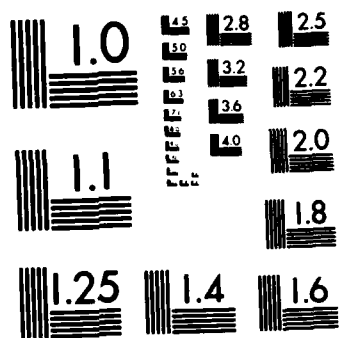
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

TABLE 2: REACTIONS OF  $\text{OH}^-(\text{H}_2\text{O})_n$  STUDIED AS A FUNCTION OF TEMPERATURE\*

Neutral Reactant	Degree of Solvation			
	n = 0	n = 1	n = 2	n = 3
(9) HF	6	6	4	3
(10) HCl	5	4	4	0
(11) HBr	6	8	7	3
(12) HI	1	1	1	1
(13) $\text{CH}_3\text{F}$	0	4	1	0
(14) $\text{CH}_3\text{Br}$	5	5	7	0
(15) $\text{CH}_3\text{Cl}$	5	4	4	0
(16) $\text{CH}_3\text{CN}$	0	4	1	0

\* Each numerical entry in this table indicates the number of temperatures in the range 200 - 500K for which rate constants were measured for a given pair of reactants.

Because of the very large amount of data collected for these reactions, only a preliminary analysis has been performed at this time. However, some representative results are included in this report.

The effect of stepwise solvation of the reactant anion  $\text{OH}^-$  upon the rates of its reactions with HBr and with  $\text{CH}_3\text{Br}$  at room temperature is illustrated in Figure 1. Increasing reactant solvation can be seen to cause only a slight decrease in the rate constant for the proton transfer reaction with HBr. By contrast, the rate of the nucleophilic displacement reaction with  $\text{CH}_3\text{Br}$  is reduced by a factor of 2 by the addition of a single  $\text{H}_2\text{O}$  molecule to the  $\text{OH}^-$ , and is further reduced by

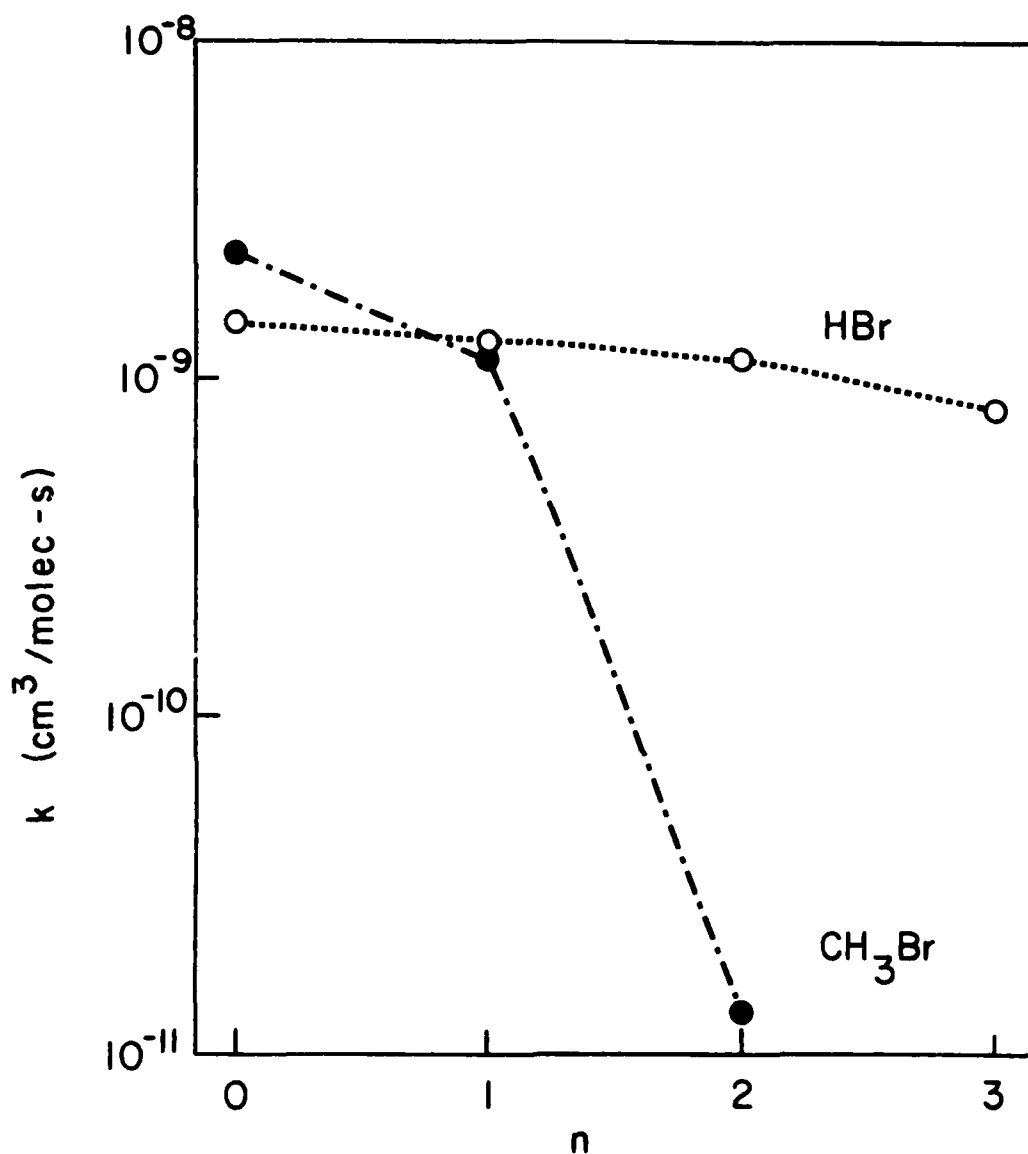


Figure 1. Effect of reactant solvation upon the rate constant  $k$  at 298 K for the reactions of  $\text{OH}^-(\text{H}_2\text{O})_n$  with HBr and  $\text{CH}_3\text{Br}$ .



a factor of 100 by the addition of a second  $\text{H}_2\text{O}$  molecule, despite the fact that formation of the unsolvated product ion,  $\text{Br}^-$ , is still exothermic by more than 14 kcal/mole. Qualitatively similar behavior is also observed for the reactions of the other hydrogen halides and methyl halides with hydrated  $\text{OH}^-$ .

The effect of temperature upon the rate constants for the reactions of  $\text{OH}^-$  with  $\text{HBr}$  and  $\text{CH}_3\text{Br}$  is illustrated in Figure 2. The rate constant for the proton transfer reaction with  $\text{HBr}$  shows little or no temperature dependence over the temperature range studied, while that for the nucleophilic displacement reaction with  $\text{CH}_3\text{Br}$  is seen to decrease with increasing temperature over this range. Again, similar behavior is observed for the reactions of the other hydrogen halides and methyl halides with  $\text{OH}^-$  and its solvates.

Although quantitative evaluation of the product branching ratios requires further data analysis, several qualitative generalizations can be made at this time. In all cases where formation of the unsolvated product anion  $\text{Y}^-$  is energetically possible, it is found to be the dominant ionic product, although small amounts (5-20%) of the singly solvated product anion were observed. In those cases where solvation of the reactant anion caused formation of the unsolvated product anion to be endothermic, the proton transfer reactions of the hydrogen halides were found to occur rapidly and with efficient transfer to the product anion of the number of  $\text{H}_2\text{O}$  molecules necessary for the reaction to be exothermic. By contrast, the rate constant for the nucleophilic displacement reactions became immeasurably small when reactant solvation caused formation of the unsolvated product anion to be endothermic.

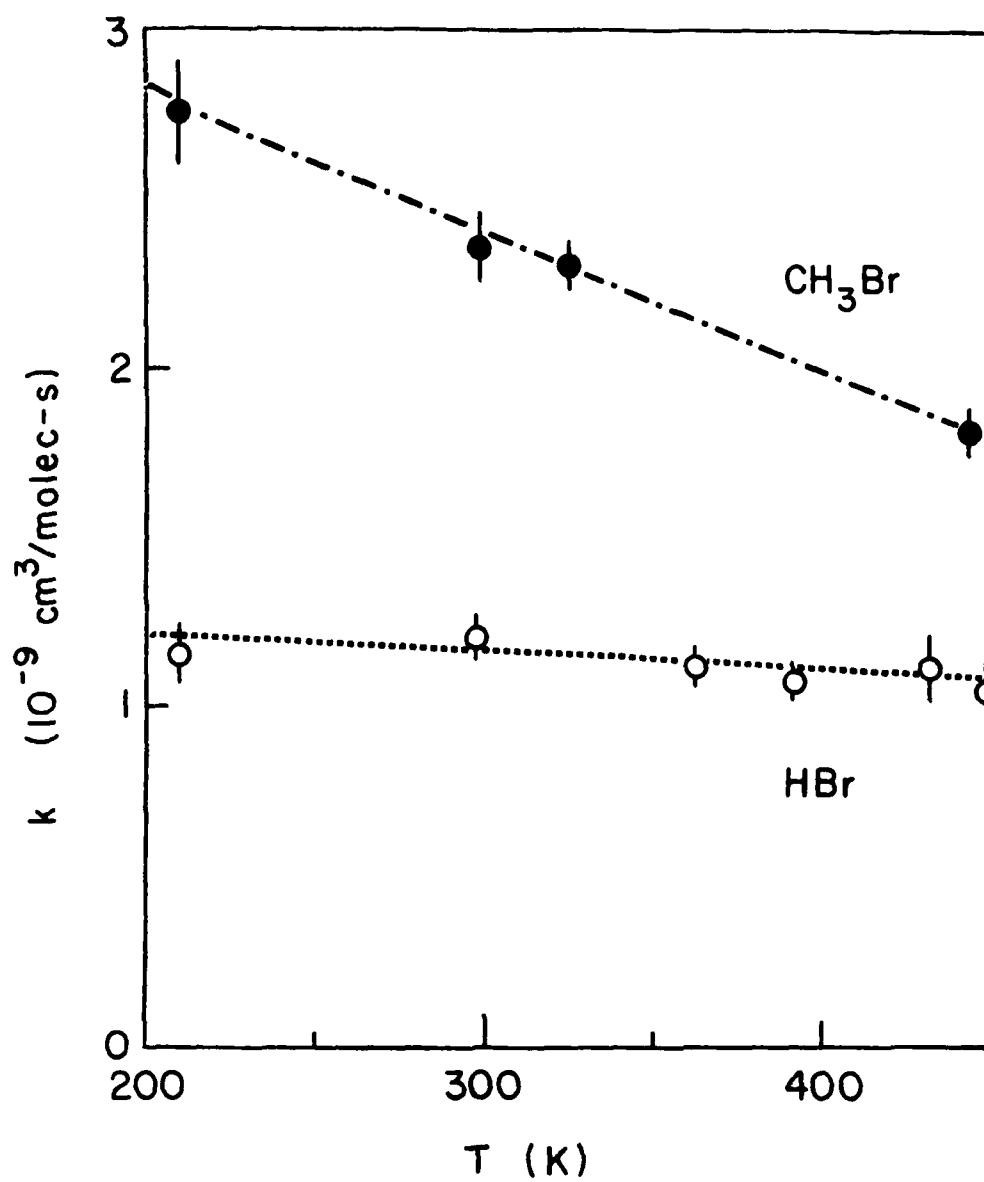


Figure 2. Effect of temperature  $T$  upon the rate constant  $k$  for the reactions of  $\text{OH}^-$  with  $\text{HBr}$  and  $\text{CH}_3\text{Br}$ .

### C. REACTIONS OF $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_n$ WITH THE METHYL HALIDES

In order to investigate the effect of solvation and temperature upon the reactions of a nucleophile other than  $\text{OH}^-$ , rate constants were measured for the reactions of the unsolvated and the singly-solvated methoxide anion,  $\text{CH}_3\text{O}^-$ , with the methyl halides as a function of temperature over the range 200 - 500K (see Table 3).

TABLE 3: REACTIONS OF  $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_n$  STUDIED AS FUNCTION OF TEMPERATURE\*

Neutral Reactant	Degree of Solvation	
	n = 0	n = 1
(17) $\text{CH}_3\text{Cl}$	6	4
(18) $\text{CH}_3\text{Br}$	7	1
(19) $\text{CH}_3\text{I}$	6	2

\*Each numerical entry in this table indicates the number of temperatures in the range 200-500°K for which rate constants were measured for a given pair of reactants.

Qualitatively, the effects of temperature and reactant solvation upon the reactions of  $\text{CH}_3\text{O}^-$  with these molecules are quite similar to those observed for the corresponding reactions of hydrated  $\text{OH}^-$ . The unsolvated  $\text{CH}_3\text{O}^-$  ion reacted at nearly the collision rate, but addition of one solvent molecule decreased the rate constant, most appreciably in the case of the least exothermic reaction. Further data analysis is required, however, before a more quantitative characterization of these reactions can be made.

#### D. MISCELLANEOUS REACTIONS AT ROOM TEMPERATURE (298K)

Rate constants and product branching ratios (where appropriate) were measured at room temperature for the reactant systems listed in Table 4:

TABLE 4: REACTANT SYSTEMS STUDIED AT ROOM TEMPERATURE (298K)

- 
- |      |                                                                      |
|------|----------------------------------------------------------------------|
| (20) | $\text{OH}^-(\text{H}_2\text{O})_{2,3} + \text{CH}_3\text{OH}$       |
| (21) | $\text{OH}^-(\text{H}_2\text{O})_{0,1} + \text{CH}_2\text{Cl}_2$     |
| (22) | $\text{OH}^-(\text{H}_2\text{O})_{0,1} + \text{CHCl}_3$              |
| (23) | $\text{OH}^-(\text{H}_2\text{O})_{0,1} + \text{CCl}_4$               |
| (24) | $\text{OH}^-(\text{H}_2\text{O})_{0,1} + \text{SiCl}_4$              |
| (25) | $\text{OH}^-(\text{H}_2\text{O}) + \text{C}_2\text{H}_2$             |
| (26) | $\text{OH}^-(\text{H}_2\text{O})_{0,1} + \text{CH}_3\text{CCH}$      |
| (27) | $\text{OH}^-(\text{H}_2\text{O}) + \text{CH}_2=\text{C}=\text{CH}_2$ |
| (28) | $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_{0,1} + \text{HF}$      |
| (29) | $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_{0,1} + \text{HCl}$     |
| (30) | $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_{0,1} + \text{HBr}$     |
| (31) | $\text{CH}_3\text{O}^-(\text{CH}_3\text{OH})_{0,1} + \text{HI}$      |
- 

Reaction (20), which involves proton transfer from the  $\text{CH}_3\text{OH}$  to the reactant anion, is of interest because simultaneous transfer of at least 2 water molecules to the  $\text{CH}_3\text{O}^-$  product anion is required for the reaction to be exothermic. We find reaction to occur on essentially every collision, indicating that such solvent transfer, although appearing to be rather complex, is actually quite facile.

Reactions (21) - (23) were studied to determine the effect of

increased substitution upon the reactivity of neutral substrate. With  $\text{CH}_3\text{Cl}$  as the neutral reactant, proton transfer was endothermic and only nucleophilic displacement was observed. With  $\text{CH}_2\text{Cl}_2$ , however, proton transfer is evidently exothermic (or at least thermoneutral), for we observe very rapid reaction to form roughly equal amounts of the proton transfer and the nucleophilic displacement products; this finding appears to be an exception to the general rule<sup>16</sup> that proton transfer, when exothermic, always pre-empts nucleophilic displacement.

Reactions (25) - (27), which also proceed via proton transfer, were found to occur with a collision efficiency considerably less than unity, and to become less efficient with increasing solvation of the reactant anion. These results are similar to those reported previously by Bohme and co-workers.<sup>9</sup>

Reactions (28) - (31) involve proton transfer to the methoxide anion from the hydrogen halides, and give results very similar to those reported above for the proton transfer reactions of these neutrals with the hydroxide anion (and its hydrated forms): reaction occurs with nearly unit efficiency and is little affected by the addition of a single solvent molecule to the reactant anion.

A more detailed characterization of these reactions requires further analysis of the data collected this summer.

## V. RECOMMENDATIONS

During the relatively short period of this project, a very large amount of data has been collected on the effects of temperature and reactant solvation upon the rates of gas-phase ion-molecule reactions. These results, although only preliminary in scope and, as yet, incompletely analyzed, have greatly increased the available body of

data on such effects and amply demonstrate the suitability of instruments such as the AFGL SIFT for such studies. Because such studies are vitally important for an improved understanding of ionic reactions, both in the gas-phase and in solution, the following recommendations are made:

(1) The data collected this summer should be analyzed and the results disseminated by publication in scientific journals. Analysis would involve both the averaging of the rate constants measured in the N replicate runs for each reaction studied, and the evaluation of the product branching ratios for those reactions in which more than one set of products are formed. Publication would involve the preparation of manuscripts in which the present results are presented and compared with any related studies previously published.

(2) The present experimentally measured rate constants should be compared with those calculated from current theoretical models. A measure of the efficiency of a given reaction can be obtained by comparing the measured rate with the collision rate, which might, for example, be calculated from the AADO theory.<sup>17</sup> The collision efficiency and its measured temperature dependence might then be compared with the predictions of a model such as the "double-minimum" model proposed by Brauman and co-workers.<sup>13,18</sup>

(3) Experimental studies of the type reported here should be continued and extended. Such future studies should not only fill in the gaps in this preliminary study, but should also be extended to include other reactant anions (such as  $F^-$ ,  $Cl^-$ ,  $Br^-$ ,  $C_2H_5O^-$ , etc.), other solvent molecules (such as the hydrogen halides, ethanol, acetone, etc.), and/or other neutral reactants.

## REFERENCES

1. For recent reviews, see (a) Ionic Processes in the Gas Phase, edited by M. A. Almoester Ferreira, Reidel Publ. Co., Dordrecht (1984); (b) Gas Phase Ion Chemistry, edited by M. T. Bowers, Academic Press, New York (1979).
2. D. K. Bohme, "Gas-Phase Studies of the Influence of Solvation on Ion Reactivity," in Ref. 1(a).
3. See, for example, F. C. Fehsenfeld and D. L. Albritton, in Atmospheric Water Vapor, edited by A. Deepak, T. D. Wilkerson, and L. H. Ruhnke, Academic Press, New York (1980), pp. 587-597.
4. G. I. Mackay and D. K. Bohme, J. Am. Chem. Soc. 100, 327 (1978).
5. D. K. Bohme and G. I. Mackay, J. Am. Chem. Soc. 103, 978 (1981).
6. K. Tanaka, G. I. Mackay, J. D. Payzant, and D. K. Bohme, Can. J. Chem. 54, 1643 (1976).
7. D. K. Bohme and A. B. Rakshit, J. Am. Chem. Soc. 106, 3447 (1984).
8. S. D. Tanner, G. I. Mackay, and D. K. Bohme, Can. J. Chem. 59, 1615 (1981).
9. D. K. Bohme, A. B. Rakshit, and G. I. Mackay, J. Am. Chem. Soc. 104, 1100 (1982).
10. M. Henchman, J. F. Paulson, and P. M. Hierl, J. Am. Chem. Soc. 105, 5509 (1983).
11. N. G. Adams and D. Smith, Int. J. Mass Spectrom. Ion Phys. 21, 349 (1976).
12. J. C. Weisshaar, T. S. Zwier, and S. R. Leone, J. Chem. Phys. 75, 4873 (1981).
13. W. N. Olmsted and J. I. Brauman, J. Am. Chem. Soc. 99, 4219 (1977).
14. D. K. Bohme and A. B. Rakshit, J. Am. Chem. Soc. 106, 3447 (1984).
15. F. C. Fehsenfeld and E. E. Ferguson, J. Chem. Phys. 61, 3181 (1973).
16. J. L. Beauchamp, in Interactions Between Ions and Molecules, edited by P. Ausloos, Plenum Press, New York (1975), p. 425 ff.

17. W. J. Chesnavich, T. Su, and M. T. Bowers, in Kinetics of Ion-Molecule Reactions, edited by P. Ausloos, Plenum Press, New York (1979), p. 31 ff.
18. W. P. Olmsted, Ph.D. thesis, Stanford University, 1977 (University Microfilms, Ann Arbor, MI #77-12679).



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

A MONTE CARLO SAMPLING OF BDR TIMES

Prepared by:	Dr. Paul C. Hoffman
Academic Rank:	Assistant Professor
Department and University:	Department of Civil Engineering Villanova University
Research Location:	Air Force Engineering and Services Center Engineering and Services Laboratory Engineering Research Division Rapid Runway Repair Branch
USAF Research	Lt. Col. Dennis W. Wiedemeier
Date:	August 10, 1984
Contract No:	F49620-82-C-0035

A MONTE CARLO SAMPLING OF BDR TIMES

by

Paul C. Hoffman

ABSTRACT

A computer simulation study of BDR times was conducted. Statistical analyses of field tests were performed and repair activities identified. Simple regression analyses were applied to formulate relationships between crater size and repair activities. Based on statistical and subjective inputs a probabilistic model was developed and a basic source program compiled. Simulations were conducted to determine both mean crater repair times and the respective standard deviations. Recommendations were proposed to further enhance the simulation modeling of repair operations.

#### Acknowledgement

This research was sponsored by the Air Force Office of Scientific Research and managed by the Southeastern Center for Electrical Engineering Education. The research was conducted at the Air Force Engineering and Services Center, Tyndall AFB, Florida. The author wishes to thank the Engineering and Services Laboratory for its hospitality and assistance.

Finally, he would like to express his appreciation to the Rapid Runway Repair Section for their cooperation and enthusiasm. In particular, special thanks is owed to Lt. Col. Dennis W. Wiedemeier for the initial problem statement and the many helpful discussions.

## I. Introduction:

Crucial to rapid runway repair (RRR) is the selection of the minimum operational strip (MOS). The "minimum" being defined as that portion of the damaged runways and taxiways that could quickly be repaired to allow aircraft launch and recovery. The selection criteria must consider not only repair duration but, also the confidence (probability) of successful completion within the required time limitation. Every alternative, as to a MOS selection, must be evaluated with respect to a time constraint and a corresponding confidence limit.

Presently, bomb damage repair (BDR) techniques are undergoing extensive evaluation. However, actual field testing is costly. Consequently, field observations must be limited to as few as possible. To augment the evaluation process mathematical models need to be developed. Then the models could be used to forecast results which in turn would provide a comparison between the various repair techniques.

One such analytical model of a repair procedure has been developed (1). The crater repair process was defined as five separate activities as follows:

1. Clear debris from crater lip and mark area to be repaired.
2. Prepare crater for fill.
3. Fill crater.
4. Compact and grade.
5. Install fiberglass mat.

From field data, functional relationships were extrapolated

for activity times versus apparent diameters. The model predicts deterministically total BDR time for a specified apparent diameter. Such a deterministic approach cannot account for the uncertainty that arises in the repair activities.

A cursory review of the observed activity times for crater repair indicated a substantial scatter in the results [2]. For a mathematical model development, it became obvious that a probabilistic approach needed to be adopted. In other words, the projection of the time to complete a bomb crater repair should include a statement as to the reliability of the prediction. Such a probabilistic approach allows for a better evaluation between alternatives. For example, a commander can make a much more informed decision if the alternate forecasted repair times, 2 and 2.5 hours, had appended to them reliabilities of 75 and 95 percent, respectively. If the mission permitted, a commander would select the larger forecasted time, 2.5 hours, in view of the greater reliability.

## II. OBJECTIVE:

A complete study of the uncertainty aspect was beyond the scope and time constraint of this project. Subsequently, this project focused on the randomness of total repair times for the fiberglass mat repair technique. The research objectives were:

1. To develop relationships between crater diameter and activity times based on simple regression analysis.

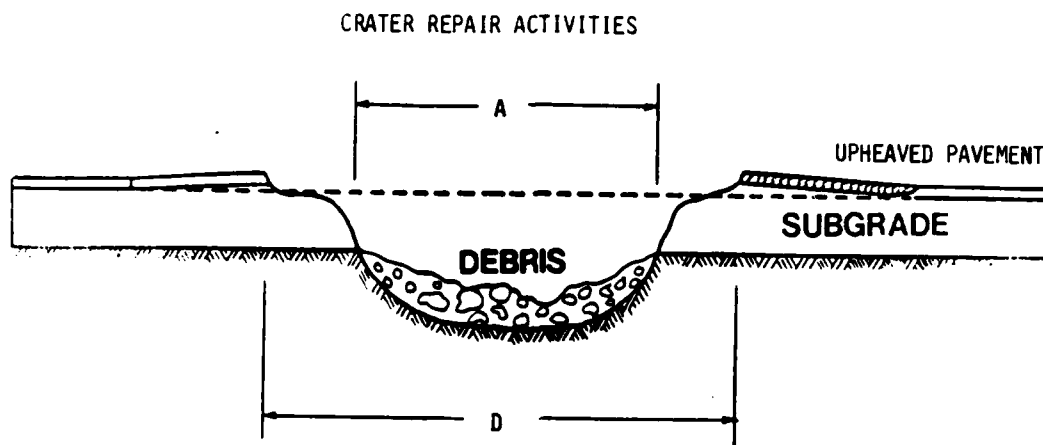
2. To incorporate the regression relationships in a Monte Carlo simulation model.
3. To evaluate the magnitude of the variations in repair times from the simulation results.
4. To postulate an decision model for RRR research.

In summary, the research project was a feasibility study of a possible scheme whereby RRR techniques could be compared. The more complex problem of combining vague information with field data was left as a follow-up study.

### III. MODEL FORMULATION:

The initial portion of the model development was the formulation of the repair diameter,  $D$ , to that of the apparent diameter,  $A$ , Figure 1. Next, possible relationships of activity times, as previously defined, versus repair diameter were investigated. The data analysis was limited to tests conducted on exploded craters [1]. Simple regression formulations were sought, Table 1. The regression relationships were based upon optimizing the coefficient of determination,  $R^2$ , while at the same time limiting the independent variables to the repair diameter and apparent diameter. As seen in Table 1, certain relationships provided reasonable results while in other cases an acceptable relationship could not be determined.

Once the best relationships were determined the probabilistic model had to be developed, Figure 2. In the four cases where the regression analysis produced relationships, the dependent variable was taken as normally distributed about the regression equation. For the other two



- T<sub>1</sub>: Clear and Mark Crater
- T<sub>2</sub>: Prepare for Fill
- T<sub>3</sub>: Fill Crater
- T<sub>4</sub>: Compact and Grade
- T<sub>5</sub>: Install Fiberglass Mat

Figure 1: Crater Profile and Activities

Table 1: Relationships By Simple Regression Analysis

Dependent Variable	Independent Variable	Relationship (standard deviation)	Coefficient of Determination
D	A	$D=0.36+1.32A$ (1.95)	$R^2=0.9595$
$T_1$	D	*	*
$T_2$	D	$T_2=-62.33+6.17D$ $-0.189(D^2-A^2)$ (4.80)	$R^2=0.9028$
$T_3$	D	$T_3=-81.01+11.54D$ $-0.42D^2+0.005D^3$ (7.14)	$R^2=0.6044$
$T_4$	D	$T_4=0.073D^2$ (12.30)	$R^2=0.5397$
$T_5$	D	*	*

\* No significant relationship was found.



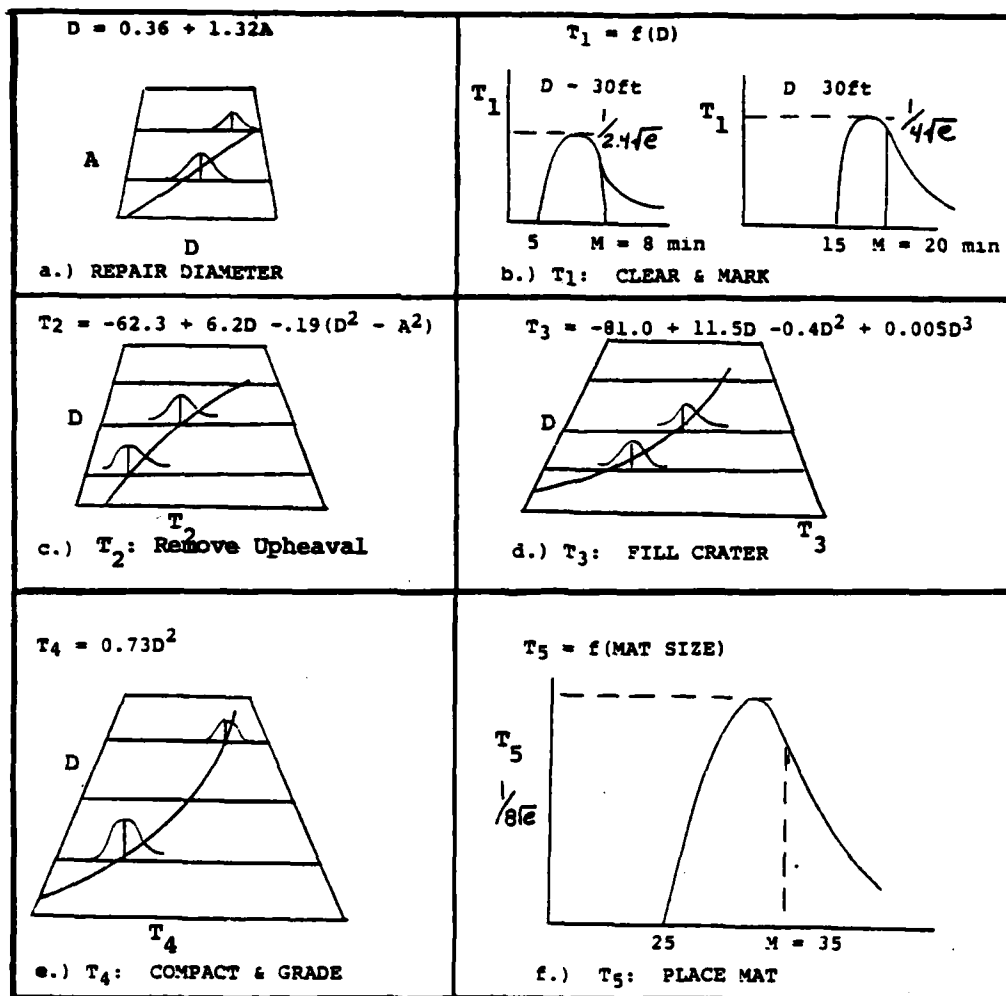


Figure 2: Probabilistic Model

cases, a Rayleigh probability density function (pdf) was adopted. For this initial feasibility study, the use of the Normal and Rayleigh pdf's seemed acceptable [3]. For the normal pdf's the variances were taken as the variances around the respective regression equations multiplied by an amplification factor. The amplification factor maintains the same confidence as one uses the equation for values away from the mean. The factor was determined as;

$$\text{Factor} = 1 + 1/n + (X_0 - \bar{X})^2 / \sum (X - \bar{X})^2 \quad (1)$$

where,  $n$  = number of data points.

$X_0$  = value of the independent variable.

$\bar{X}$  = mean of the independent variable.

$X$  = observed values of the independent variable.

#### IV. COMPUTER ALGORITHM:

Within most computer packages a random number generator exists which produces random numbers over the interval [0,1]. The basis of digital simulation is the transformation of a nonuniform random distribution to a computer generated uniform random distribution. If  $U$  is a uniformly distributed random number over the interval [0,1], then it follows that;

$$P(U \leq u_1) = F_U(u_1) = u_1 \quad (2)$$

For a random number  $X$  with a nonuniform distribution, there must exist an  $x_1$  such that,

$$F_X(x_1) = u_1 \quad (3)$$

and so,

$$P(X \leq x_1) = P(U \leq u_1) \quad (4)$$

In other words, for computer simulation the value of  $u$  is the area under the pdf,  $f_X(x)$ , of the random number  $X$ . Formally, this is defined as the inverse function,

$$F_X^{-1}(u_1) = x_1 \quad (5)$$

For a Rayleigh density function,

$$f_X(x) = \begin{cases} 0 & x < 0 \\ (x^2/a^2)e^{-x^2/2a^2} & x \geq 0 \end{cases} \quad (6)$$

where  $a^2$  is the density parameter, the inverse function is analytically determined as;

$$x = [-2a^2 \ln(1 - F_X(x))]^{1/2} \quad (7)$$

For a standard normal density function, an analytical inverse function is intractable, but an approximation can be made as follows [4];

$$F_X(x) = 1 - r(a_1 t + a_2 t^2 + a_3 t^3) \quad (8)$$

where  $t = (1 + 0.3326x)^{-1}$  and  $a_1$ ,  $a_2$ ,  $a_3$ , and  $r$  are known constants. The solution, for a given  $F_X(x)$ , can be found by an iterative method.

The Monte Carlo Model samples from a repetitive generation of random numbers, Figure 3. For a specified apparent diameter the mean repair diameter is computed from the regression relationship. A random number is generated and the actual sample repair diameter is determined iteratively. For the computed repair diameter the mean activity times are calculated. From sequential generation of random numbers and

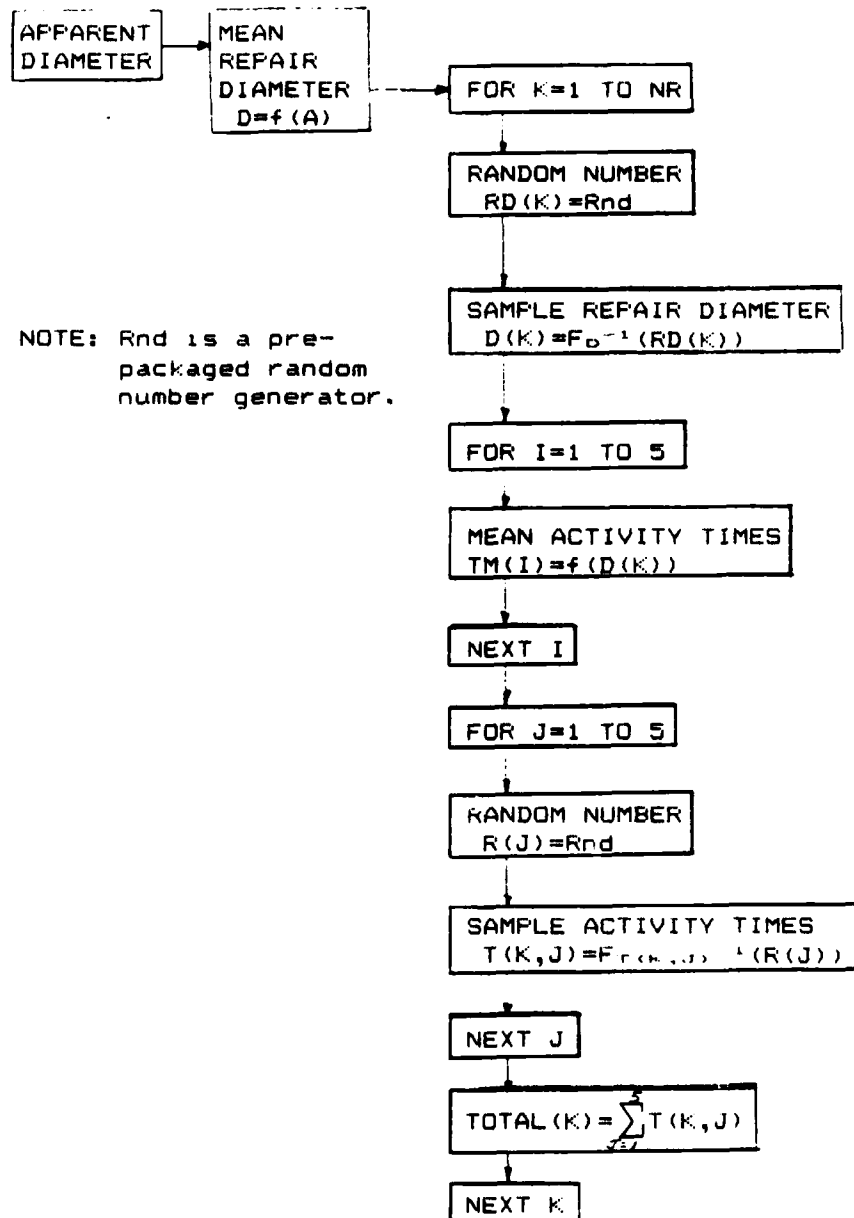


Figure 3: Monte Carlo Model

application of the appropriate inverse functions the activity times are simulated. The summation of the activity times is a sampled crater repair time. This process is repeated for the specified number of simulated trials, NR.

#### V. MONTE CARLO SAMPLES:

The results of the simulation study have been summarized, Table 2. For comparative purposes, experiments of 200, 400 and 1000 simulations were conducted for the apparent diameters indicated. For the simulations of 1000 trials histograms were developed, Figure 4. As indicated by the histograms, the total repair times seem to be normally distributed. This was expected based upon the Central Limit Theorem in Statistics and the fact that four of the six density functions in the model were assumed as normally distributed. If the activity times and repair diameter had varied density functions or large differences in variances the outcome may differ markedly from a normal pdf.

Since the outcome implies a normal distribution a traditional approach with regard to confidence level and precision may be conducted, [ 5,Chapter 7]. Such a statistical analysis indicates that, based on a simulation of 1000 trials, the probability is better than 99% that the sample mean is within two minutes of the population mean. Likewise, for 1000 trials, a confidence level of 98% exists that the sample variance is within 10% of the population variance.

The actual significance of the findings is seen in the computed standard deviations. First, an extra resource that

Table 2: Averaged Results Of The Monte Carlo Tests

App Dia.  (feet)	No. of Tests	Number Of Samples							
		200		No. of Tests	400		No. of Tests	1000	
		Mean	Std. Dev.		Mean	Std. Dev.		Mean	Std. Dev.
		(min)	(min)		(min)	(min)		(min)	(min)
10	2	72.4	21.9	0	*	*	0	*	*
15	2	119.9	16.9	0	*	*	1	120.8	16.6
20	6	160.9	16.9	1	161.0	16.5	1	160.6	17.0
25	3	209.0	16.9	0	*	*	1	208.7	17.3
30	4	252.5	17.9	1	253.4	18.0	1	252.5	18.1
35	2	311.9	18.7	1	309.9	19.6	1	311.6	19.5
40	3	394.0	22.6	1	395.5	23.6	1	393.2	22.0

\* A Monte Carlo Test was not conducted.

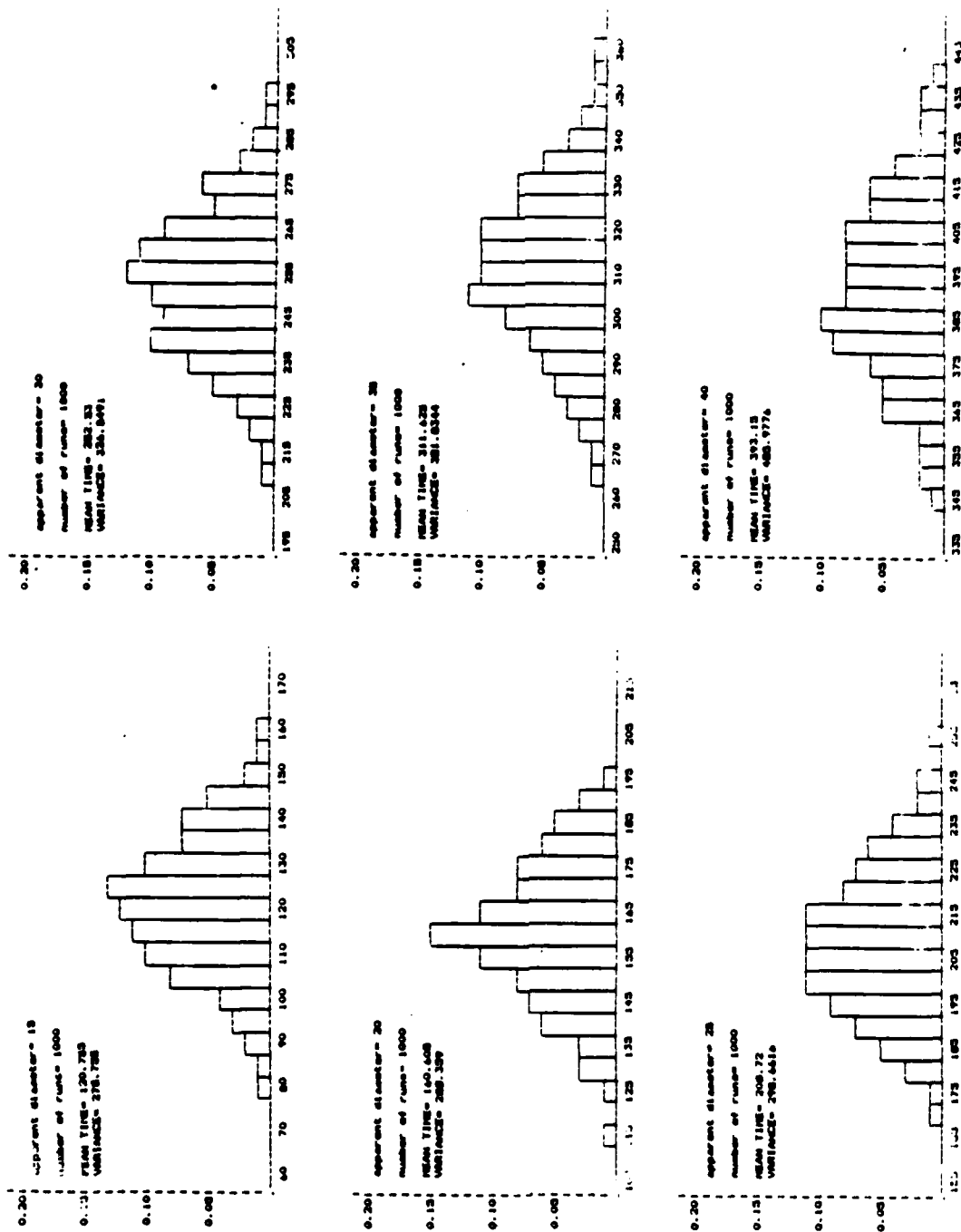


Figure 4: Monte Carlo Results

reduces the overall repair time by approximately ten minutes is insignificant since the standard deviations are expected to range from 16.9 to 22.6 minutes. It would be more prudent to commit resources for the reduction of the standard deviations rather than for ten minute reductions on the means. Second, for a 90% reliability the serial repair of two fifteen foot apparent diameter craters would require 272.1 minutes. In contrast, a 90% reliability for the repair of one thirty foot apparent diameter crater totals 276.0 minutes. In other words, for a high reliability the difference between two small craters and one large crater seems insignificant. It should be noted that these results are based solely on the field test data. Better methods, mechanization and experience will surely alter these findings.

#### VI. RECOMMENDATIONS:

The study did indicate that a simulation approach is feasible. However, during the research effort shortcomings became obvious. For example, in 200 trials for a ten foot apparent diameter crater one sample resulted in a negative completion time. Even though the model indicated that such an occurrence would be extremely rare, it is impossible from a practical point of view.

For a more realistic simulation, it is recommended that four points be incorporated as follows:

1. Modified probability density functions.
2. Refined activity definition.
3. Bayesian statistics.
4. Fuzzy set theory.



To eliminate the possibility of the model generating a negative or extremely small repair time several modified density functions should be considered. For one approach, the normal density functions could be truncated at some minimum time. Another approach could be the adoption of alternate density functions such as the log normal or gamma density function.

For a more realistic model some of the initial five activities need refinement. For example, the third activity (fill crater) should be defined as three sub-activities. First, there should be a density function for the loader operation at the stockpile. Second, the hauling operation to the crater needs to be modeled. Third, a density function for the queuing process at the crater as a function of available equipment should be developed.

With the formulation of a refined model an extensive study of crater repair should be undertaken. With a large sample size from a computer simulation, the maximum-likelihood method could be applied to more accurately estimate the parameters for the probability density function of total repair times. Once a model for total repair times is developed, updating with additional testing can be achieved with the application of Bayesian Statistics.

Finally, the question as to the incorporation of vague information needs to be answered. In any research and development project several variables cannot be observed in a laboratory environment. The effect of weather on repair times cannot be realistically tested in the field. The routing of

the trucks to the crater cannot be known precisely. Such variables might be best handled linguistically. For example, weather can be stated as being "good" or "bad". Also, the trucks might be evaluated as having a "long", "medium" or "short" haul distance. The concept of linguistic variables is known as fuzzy data. Therefore, the applicability of fuzzy set theory to the RRR research program should be investigated.

## REFERENCES

1. Wiedemeier, D. W., Time Estimates For Interim Bomb Damage Repair Techniques (Draft Report), AFESC, 1984
2. Hoffman P. C. , AFESC/RDCR Visit, March 1984
3. Benjamin, J. R. and Cornell, C. A., Probability, Statistics, and Decision for Civil Engineers, McGraw Hill Book Company, New York, 1970
4. Poole, L., Borchers, M., and Burke, P. M., Some Common Basic Programs: IBM Personal Computer Edition, McGraw-Hill, Berkeley, 1982
5. Mize, J. H. and Cox, J. G., Essentials of Simulation, Prentice Hall, Englewood Cliffs, 1968.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

BISMUTH SILICON OXIDE: SAMPLE VARIABILITY

INVESTIGATED WITH THERMALLY STIMULATED

CONDUCTIVITY AND THERMOLUMINESCENCE

Prepared by:	Brian W. Holmes
Academic Rank:	Assistant Professor
Department and	Department of Physics,
University:	San Jose State University
USAF Research:	Charles Woods
Date:	September 29, 1984
Contract no:	F49620-82-C-0035

BISMUTH SILICON OXIDE: SAMPLE VARIABILITY

INVESTIGATED WITH THERMALLY STIMULATED

CONDUCTIVITY AND THERMOLUMINESCENCE

by

Brian W. Holmes

ABSTRACT

We used thermally activated conductivity (TSC) and thermoluminescence (TL) to survey a number of different crystals of bismuth silicon oxide,  $\text{Bi}_{12}\text{SiO}_{20}$  (BSO). Samples were irradiated at liquid nitrogen temperatures with uv light from anHg lamp and then warmed to room temperature. Three principal electron traps are observed, resulting in peaks in TSC and TL; the peak temperatures ( $T_m$ ) were near 160 K, 235 K, and 298 K. Observed TSC and TL signals varied from crystal to crystal, but were similar for different samples of the same crystal. However, the TSC and TL spectra of a crystal differ. We offer a simple explanation of the systematic differences in peak heights and  $T_m$  values that are seen in comparing TSC and TL results in individual samples. Using initial rise techniques, we obtain thermal activation energies ( $E_A$ ) consistent with an earlier measurement  $E_A = 26 T_m$ . Attempts to calculate trap densities yield results which differ by several orders of magnitude. We propose further work to learn whether the variations in TSC and TL spectra are correlated with variations in trap densities, and to gather stronger evidence of the fundamental nature of electron traps in BSO.

## I - INTRODUCTION

Bismuth silicon oxide (BSO) has the chemical formula  $\text{Bi}_{12}\text{SiO}_{20}$ . This material is capable of reversible, real-time recording of holograms with a thousand times the photosensitivity of traditional silver halide emulsions.<sup>1</sup> BSO has great potential in optical data processing, but optical characteristics vary significantly, even among crystals identically prepared.<sup>2</sup> The production of uniform or improved BSO crystals may well depend on achieving a better understanding of the fundamental processes responsible for its photosensitivity.

In BSO and other photorefractive materials such as bismuth germanium oxide (BGO), absorbed photons cause electrons to diffuse through the crystal or to drift under an applied electric field. Electrons tend to be trapped in regions of weaker illumination, and the resulting uneven charge distribution causes variations in the refractive index, due to the electro-optic effect. The variations in refractive index mimic the spatial intensity variations of the incident light, so that optical information storage has occurred. An important point is that the centers responsible for electron trapping have not been identified.<sup>2</sup>

The two methods for studying electron trapping in solids are thermally stimulated conductivity (TSC) and thermoluminescence (TL).<sup>3</sup> These techniques entail irradiation of samples at low temperature, filling the electron traps, followed by warming. During warming the electrons are released from the traps, contributing to the conductivity when they are promoted to the conduction band, and contributing to luminescence if they lose energy via a luminescent center. These techniques yield information about trap energy levels and they do not call for elaborate instrumentation. However, these

techniques need to be supplemented by other studies in order to yield information on the specific nature of the observed traps."

So we chose to study trapping in BSO through TSC and TL, surveying a number of samples in order to assess these techniques in evaluating sample variations. These methods have been used earlier in Lauer's study of BSO,<sup>5</sup> but they have not previously been used in the sort of survey undertaken here.

Section II of this report states the objectives of this research effort. Section III summarizes evidence concerning electron traps and electron energy levels in BSO. Section IV discusses the experimental methods used. Section V discusses the experimental results. Section VI gives recommendations for future work on BSO.

## II- OBJECTIVES

1. To develop instrumentation for performing TSC and TL experiments on BSO.
2. To survey a number of different BSO crystals, and to see what information TSC and TL can provide on sample variability.

## III - ELECTRON ENERGIES IN BSO

Experiments at liquid nitrogen temperatures (room temperature) show a large increase in photoconductivity for excitation energies of 3.4 ev (3.25 ev). This corresponds to the energy gap between the valence and conduction bands. However, there is a shoulder in the optical absorption at excitation energies of 2.80 ev (2.60 ev), implying the existence of a state 2.80 (2.60) ev below the conduction band, normally occupied by electrons. (Electrons are shown to be the principal carrier in pulsed uv experiments.) This state is attributed to silicon vacancies in the lattice, having a density of  $10^{19} \text{ cm}^{-3}$ ; these vacancies have been inferred from x-ray diffraction studies of BGO,<sup>7</sup> but they have not been directly studied in BSO.

A sample at liquid nitrogen temperatures undergoing excitation at the band gap shows visible red luminescence centered at 1.95 ev, quenched as the temperature rises. This luminescence does not involve free charge, so the center responsible for it would not appear on an energy-band diagram.<sup>8</sup> This luminescence has been attributed to  $\text{Bi}^{3+}$ .<sup>9</sup>

Another luminescence, centered at 1.3 ev, is seen in thermoluminescence. Here free charge is involved, since both the TL and photoconductivity are enhanced by ir irradiation. The TL is excited by irradiation at the band gap and by irradiation peaked at 2.25 ev. The center responsible for this luminescence is 1.3 ev below the conduction band when occupied by a hole, and 2.25 ev below when occupied by an electron.

TSC and TL results<sup>5</sup> show that electrons are trapped 0.34, 0.54, and 0.65 ev below the conduction band. The density of these traps is  $10^{15} \text{ cm}^{-3}$ , smaller than the density of the traps ( $10^{16} \text{ cm}^{-3}$ ) responsible for the photorefractivity.<sup>1</sup> Energies smaller than 0.17 ev and larger than 0.8 ev have not been covered by TSC and TL measurements.

Figure 1 gives an energy-band diagram for BSO based on the preceding considerations.

#### IV - EXPERIMENTAL DETAILS

BSO has a cubic structure;<sup>8</sup> crystals are grown using the Czochralski technique. The starting material, mostly bismuth oxide, is highly reactive at crystallization temperatures. Crystals are grown in gold crucibles and may contain small amounts of dissolved gold.<sup>10</sup>

Itek provided a number of samples. Growth runs were S, G, Q, and J, sequentially. Growth methods and starting materials have not varied since the middle of the G-runs. Sample G-17-18 is the 17th crystal in the G-sequence, and the sample is the 18th piece cut from the boule. Good samples show low



birefringent strain; but most of the samples provided are rejects. No aluminum-doped samples were available.<sup>6</sup> Many samples showed a tendency to fracture under heating and ultrasonic soldering, especially the samples provided by Crystal Technology.

Samples were cut in  $0.5 \times 0.5 \text{ cm}^2$  and  $1.0 \times 1.0 \text{ cm}^2$  sizes. The thicknesses varied from 0.08 to 0.17 cm. Occasionally the edges were polished. Electrical contact was made through ultrasonic indium soldering.<sup>5</sup> This produced cracks in some of the samples; evidence indicates that the cracks often reflect strain already present in the crystals. Contacts were ohmic in the range 5-45 volts; this determination is difficult, owing to the high dark resistivity ( $10^{15} \text{ ohm cm}$ ) of BSO.

Samples were pressed against a copper mount inside an optically accessible dewar. The mount was in thermal contact with a copper rod connected to a liquid nitrogen reservoir and a heater. When voltage was applied to the heater from a variable voltage power supply, the liquid nitrogen evaporated and the temperature of the mount rose. The temperature of the copper mount was monitored with a copper-constant ~~ine~~ thermocouple pressed into a hole in the mount. The electrical leads of the thermocouple were varnished to the mount to reduce heat leaks; however, the varnish did not thrive under the thermal cycling of the apparatus.

Thin Teflon plumbing tape isolated the samples electrically from the copper mount. The samples were pressed into the Teflon tape by Teflon fasteners screwed to the copper mount. Mica was also tried, but the Teflon tape had better thermal conduction.

Samples were irradiated at liquid nitrogen temperatures using unfiltered light from a mercury lamp. The lamp was warmed up four minutes prior to use

to ensure constant intensity. Attempts were made at filtering the irradiating light, but these efforts were not systematic.

Irradiation was usually for ten minutes. This gave TSC spectra which were not significantly different from the spectra resulting from a twenty-minute exposure. The spectra did not change if the sample was electrically biased during irradiation. Biasing enabled recording the photocurrent, which tended to decrease as a function of time. The time behavior of the photocurrent was not systematically investigated.

After irradiation, the dewar windows were covered and the heater was activated. Typically the heating rate was 0.8 K/sec. TSC was measured using a Keithley 103B electrometer. TL was measured using an RCA 7102 phototube with an ir sensitive S-1 cathode. The tube was cooled by nitrogen gas drawn from a liquid nitrogen dewar; this reduces the dark current significantly. Dewing of the tube window was eliminated by flowing air through a collar surrounding the window. The output of the tube was subject to shifts, due to temperature instabilities and background light, and spikes, possibly due to electrical charge buildup and discharge on the tube surface. The signal strength was basically good, but the cooling system was troublesome.

Data was collected on x-y recorders, giving the TSC or TL as a function of thermocouple voltage.

#### V - DISCUSSION

A number of different irradiation times was tried. Figure 2 gives a sequence of TSC measurements for 2.5, 5, 10, and 20 minute irradiations. There is no significant difference between the last two; most measurements reported here were for samples exposed 10 minutes.

In a few cases, TSC measurements were made for a number of different bias voltages. Increased voltage increased the current, but there were only slight changes in the shapes and positions of the TSC curves: see Fig. 3.

Irradiated samples were observed to darken visibly. The darkening disappears as the sample warms. This effect, which does not seem to be described in the literature, was not investigated systematically.

Figure 4 gives TSC and TL for a representative sample. In general, the TSC and TL spectra show three major peaks as the temperature rises. We number these peaks 1, 2, and 3 in order of increasing temperature. Table 1 gives the temperatures  $T_m$  associated with the maxima for each sample. The peaks occur near 160 K, 235 K, and 298 K. These values are reasonably close to those seen in Figure 1 of Lauer's paper.<sup>6</sup> In general, our spectra are similar to Lauer's, except that we do not see as much detail in TSC for temperatures greater than 300 K. However, Lauer's TSC and TL spectra have about the same relative peak sizes; we do not get this agreement. Peak sizes will be discussed below.

Considering the peak temperatures in detail, we note that peak 1 in TSC seems to be a double peak; this feature is strongly in evidence in the results shown in Fig. 3. This double peak does not yield a consistent  $T_m$ . Neglecting double peaks, TSC and TL agree well for  $T_m$  of peak 1, except for sample Q-2-22. For peak 2, the TL peaks at lower temperatures than the TSC; the average lead is 1.8 K. For peak 3, TL leads TSC by an average of 6.2 K. We will explain these temperature shifts below.

Peak heights depend on the heating rate, applied voltages, and sample thicknesses for TSC; they depend on heating rate, sample size, phototube voltage, and experimental geometry for TL. Table 2 gives the peak heights for each sample, normalized so that the second peak has a height of unity. A

consistent pattern emerges: the height of peak 3 is always greater than the height of peak 2 for TSC, and the height of peak 3 is always less than the height of peak 2 for TL.

Consideration of the photoluminescence emission at 1.30 eV, given in Fig. 5 (modified from Fig. 7 of the paper of Hou *et al.*<sup>6</sup>) offers useful insight in understanding the relative peak heights and  $T_m$  shifts of the TL compared with the TSC. The photoluminescence occurs under illumination at the band-gap energy. This promotes electrons into the conduction band, from which they may return to the valence band by way of the 1.3 eV luminescence center. In our case, the electrons are provided not by UV irradiation but by thermal activation from shallow electron traps. The photoluminescence is quenched at higher temperatures; this explains why peak 3 for TL is small compared with peak 3 for TSC. This means that the electrons have more non-radiative ways of losing energy at higher temperatures. The quenching also explains why this peak is shifted to lower  $T_m$  values for TL compared with TSC. We note that peak 2 is large for TL compared with TSC; this is consistent with the peak in photoluminescence emission at 212 K, near peak 2 at 235 K. The slight shift of  $T_m$  to lower temperatures in peak 2 of TL compared to TSC is related to the increase in photoluminescence intensity for temperature below 235 K. Peak 1 is near a local minimum in the photoluminescence emission, so this peak in TL is diminished in strength in comparison with TSC.

In most cases, each BSO sample was available in 0.5 x 0.5 cm<sup>2</sup> and 1.0 x 1.0 cm<sup>2</sup> sizes. Table 3 gives normalized heights of peak 1 and peak 3 for TSC in large and small specimens of each sample. This disagreement is worst for sample G-27-6, and the relative height of peak 3 seems larger in small samples than in large samples. In general, the agreement between small and large specimens is satisfactory; peak shapes and  $T_m$ 's also compare well. Figure 6

gives TSC spectra of small and large specimens of G-25-11. We note also that BSO crystals are not isotropic in strain or impurities<sup>10</sup>, which means small and large specimens could be different enough to yield different TSC spectra.

Table 4 gives calculated thermal activation energies of the three electron traps, expressed as multiples of  $kT_m$ . These were calculated using 'initial rise' techniques.<sup>5,14</sup> Lauer<sup>6</sup> obtained  $E_A = 26T_m$ , based on similar analysis of his results. Our results are consistent with this, except for peak 1. If peak 1 is double, then there is some ambiguity in  $T_m$ . Averaging over the results in Table 4, we find the activation energies of peaks 2 and 3:

$$E_2 = (26.7 \pm 2.3)kT_m \text{ (TSC)} = (25.5 \pm 3.8)kT_m \text{ (TL)}$$

$$E_3 = (26.3 \pm 1.8)kT_m \text{ (TSC)} = (24.5 \pm 1.8)kT_m \text{ (TL)}.$$

It is difficult to calculate the density of electron traps from the TSC and TL data. This is because it is not known (from these experiments) how long an electron remains in the conduction band. Using the known heating rate (0.8 K/sec), one can integrate under a peak in the TSC spectrum to estimate the trap density, assuming each electron represents one trap. For peak 1 of Q-17-18, the density thus obtained is  $4 \times 10^{12} \text{ cm}^{-3}$ . Hou *et al.*<sup>6</sup> calculate a drift range of  $8.5 \times 10^{-7} \text{ cm}^2/\text{V}$ ; our field was 10 volt/cm, giving a drift of  $8.5 \times 10^{-6} \text{ cm}$ , much smaller than the sample width. Hence the actual trap density could be much larger. For TL in a similar calculation gives a trap density of  $4 \times 10^{14} \text{ cm}^{-3}$ . This is not far from the value  $10^{15} \text{ cm}^{-3}$  reported by Lauer<sup>16</sup> for peak 3, but several assumptions involved in this calculation are unrealistic. The discrepancies between these two calculations of trap densities and the dubiousness of some of the assumptions on which the calculations are based emphasize the need for using other experimental techniques to supplement TSC and TL in calculating trap densities.

To summarize, this study has shown that TSC and TL spectra of BSO crystals vary distinctively from sample to sample. The measurements are easy to make, and they are easy to extend and improve. Although there are three principal electron traps responsible for the results described, the relative contributions of the traps vary from sample to sample. On the other hand, the peak temperatures of the peaks do not vary much. These traps probably do not participate in the photorefractivity of BSO; a similar survey, of properties relevant to four wave mixing, is in order.<sup>2</sup> It will be interesting if the variations observed in such a study correlate with the variations reported here.

#### VI - SUGGESTIONS FOR FURTHER WORK

1. Irradiation with light of different wavelengths should be undertaken, yielding excitation spectra for TSC and TL.
2. The emission spectrum for the TL should be measured.
3. Data collection should be automated; this would greatly simplify the analysis.
4. Absorption spectra should be obtained for these samples, preferably as a function of temperature. These measurements may give useful indications of the density of trapping centers.<sup>2</sup>
5. The temperature range should be extended to higher and lower temperatures; this will reveal deeper or shallower electron traps, which have been predicted.<sup>6</sup>
6. X-ray diffraction studies should be performed on BSO to verify the existence of silicon vacancies, and chemical analysis to characterize chemical impurities.
7. EPR studies may clarify the nature of the trapping centers.<sup>2,3</sup>

8. Nuclear quadrupole resonance studies have already been performed on BSO.<sup>11</sup> These measurements yield information about the surroundings of the Bi nuclei. If the Bi atoms participate in the trapping, producing a change in the electric field gradient seen by enough Bi nuclei, the changes in the NQR signal should be detectable.<sup>12</sup>

9. Suggestions have been made<sup>2</sup> that transient capacitance spectroscopy techniques (such as DLTS) might be of value. However, these techniques are most often used in diode junctions,<sup>13</sup> for which forward biasing will produce enough electrons to populate traps. This is not possible in BSO. Electrons for non-equilibrium studies might be provided by a pulsed irradiation source.

10. Doped samples might reveal information about the nature of traps. In particular, doping with a donor (such as phosphorus)<sup>6</sup> might suppress the 1.3 eV thermoluminescence by raising the Fermi level high enough to cause the trap associated with the thermoluminescence to be filled with electrons. Samples doped with iron or chromium ions would be particularly suitable for EPR spectroscopy.<sup>2</sup>

11. In view of the x-ray applications of these crystals, x-ray irradiations as well as other forms of non-uv irradiation should be undertaken.

12. Systematic studies of four-wave mixing should be undertaken and correlated with studies like the present to see what relation there is between the traps studied here and those responsible for BSO's photorefractivity.

#### ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, RADC/Hanscom AFB, and the Southeastern Center for Electrical Engineering Education for the welcome chance to do interesting physics at Hanscom. I would particularly like to thank C. Woods, C. Pitha, C.

Leiby, and J. Ludman for their help, support and interest. Others at RADC rendering valuable help were B. Ahern, E. Davies, P. Drevinsky, W. Ewing, P. Hemmer, J. Horner, H. Lipsom, T. Ryan, and J. Villieux. Valuable discussions and help came from R. Lauer of GTE-Sylvania and C. Warde of MIT. The interest, advice, and numerous samples offered by R. Aldrich of Itek were of crucial importance to the execution of this effort.



#### REFERENCES

1. M. Peltier & F. Micheron, "Volume Hologram Recording and Charge Transfer Process in  $\text{Bi}_{12}\text{SiO}_{20}$  and  $\text{Bi}_{12}\text{GeO}_{20}$ ", J. App. Phys. 48(9), 1984, pp. 3683-90.
2. Y. Ja, "The Significance of Trapping Center Concentration in Photorefractive Crystals used for Degenerate Four-Wave Mixing", Opt. & Quant. Elec. 16, 1984, pp. 355-8.
3. P. Bräunlich (ed.), Thermally Stimulated Relaxation in Solids, (Springer-Verlag, New York, 1979).
4. M. Boehm, B. Cord, A. Hofstaetter, A. Scharmann, & P. Parot, "EPR, Thermoluminescence, and Thermally Stimulated Conductivity in  $\text{BaWO}_4$  Crystal ", J. Lumin. 17, 1978, pp. 291-300.
5. R. Lauer, "Thermally stimulated currents and luminescence in  $\text{Bi}_{12}\text{SiO}_{20}$  and  $\text{Bi}_{12}\text{GeO}_{20}$ ", App. Phys. Lett. 42(5), 1971, pp. 2147-9.
6. S. Hou, R. Lauer & R. Aldrich, "Transport processes of photo induced carriers in  $\text{Bi}_{12}\text{SiO}_{20}$ ", J. App. Phys 44(6), 1973, pp.2652-8
7. R. Lauer, "Photoluminescence in  $\text{Bi}_{12}\text{SiO}_{20}$  and  $\text{Bi}_{12}\text{GeO}_{20}$ ", App. Phys. Lett. 17(4), 1970, pp. 178-9.
8. S. Abrahams, P. Jamieson & J. Bernstein, "Crystal Structure of Piezoelectric Bismuth Germanium Oxide  $\text{Bi}_{12}\text{GeO}_{20}$ ", J. Chem. Phys. 47(10), 1967, pp.4034-41.
9. M. Weber & R. Monchamp, "Luminescence of  $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ : Spectral and Decay Properties", J. App. Phys. 44(12), 1973, pp. 5495-9.
10. R. Aldrich, private communication.
11. Y. Buslaev, E. Kravchenko, V. Pachomov, V. Skorikov & G. Semin, "NQR- $^{209}\text{Bi}$  Spectra of Several  $\text{Bi}_n\text{Al}_{1-n}\text{VO}_m$ -type Crystals", Chem. Phys. Lett. 3(6), 1969, pp. 455-6.

12. R. Marino, private communication.
13. J. Ludman, Defect Levels in Neutron-Irradiation GaAs Schottky Diodes and Laser Diode Degradation (AFCRL-TR-73-0344, Physical Sciences Research Papers, No. 550, 1973).
14. R. Chen, "On the Calculation of Activation Energies and Frequency factors from Glow curves", J. App. Phys. 40(2), 1969, pp. 570-85.
15. RCA-7102 Photomultiplier Tube (Manual published by RCA).
16. R. Lauer, "Electron Effective Mass and Conduction-Band Effective Density of States in  $\text{Bi}_{12}\text{SiO}_{20}$ ", J. App. Phys. 45(4), 1974, pp. 1794-7.

Sample	Peak 1		Peak 2		Peak 3	
S-31-86	147.8*	(166.0)	234.4	(233.6)	297.8	(291.0)
G-9	165.8	(165.2)	236.5	(234.1)	299	(293.0)
G-21	163	(163.7)	239	(236.7)	300	(295)
G-25-11	143.8*	(155.6)	238.6	(236.6)	302.8	(296.0)
G-27-6	128.4*	(154.3*)	238.6	(235.5)	302.8	(294.0)
Q-2-22	159.3	(163.7)	241.0	(240.0)	304.5	(298.9)
Q-17-18	163.0	(163.9)	237.2	(234.7)	298.4	(291.7)
J-2	160.3	(165.2)	236.3	(236.1)	299.6	(296)
Cryst. Tech.	158.9		237.7		302.8	

Table 1 - Peak temperatures in K for different BSO crystals for TSC (TL). The peaks marked with an asterisk (\*) are double peaks. The S, J, and Crystal Technology samples are 0.5 x 0.5 cm<sup>2</sup>; the others are 1.0 x 1.0 cm<sup>2</sup>.

Sample	TSC		TL	
	Peak 1	Peak 3	Peak 1	Peak 3
S-31-86	0.43*	1.43	1.26	0.5
G-9	6.07	2.84	1.86	0.93
G-21	1.96	1.86	0.72	0.74
G-25-11	0.98*	2.01	<0.41	<0.66
G-27-6	0.75*	1.56	0.33*	0.49
Q-2-22	1.82	2.22	0.66	0.77
Q-17-18	3.59	1.95	1.35	0.76
J-2	2.17	2.20	1.75	0.96
Cryst. Tech.	1.72	2.08	--	--

Table 2 - Normalized peak heights (height of peak 2 = 1.0) for different BSO crystals.

Sample	Peak 1	Peak 3
G-9 Large	6.07	2.84
G-9 Small	5.91	3.74
G-17-18 Large	3.59	1.95
G-17-18 Small	4.5	2.7
G-21 Large	1.96	1.86
G-21 Small	2.07	2.16
G-25-11 Large	0.98	2.01
G-25-11 Small	0.95	2.25
G-27-6 Large	0.75	1.56
G-27-6 Small	1.08	2.62
Q-2-22 Large	1.82	2.22
Q-2-22 Small	1.72	2.56

Table 3 - Normalized peak height (height of peak 2 = 1.0) for different BSO crystals in TSC, comparing the results from large and small samples of the same crystal.

Sample	Peak 1	Peak 2	Peak 3
S-31-86	16.6*	30.2	25.0
G-9	15.8 (11.3)	23.5 (21.3)	25.3 (25.6)
G-21	18.2 (13.1)	27.1 (22.7)	27.1 (25.9)
G-25-11	22.8* (17.3)	24.3 (29.3)	27.0 (26.5)
G-27-6	26.6*	27.9 (29.5)	23.9 (24.3)
Q-2-22	21.4 (14.4)	25.1 (28.2)	29.3 (23.3)
Q-17-18	25.2 (12.2)	28.5 (22.2)	24.3 (21.5)
J-2	19.5	24.6	28.2
Cryst. Tecn.	16.5	28.7	26.6

Table 4 - Activation energies given as multiples of the peak temperature  $T_m$  for different BSO crystals for TSC (TL). Values marked with an asterisk (\*) are for double peaks.

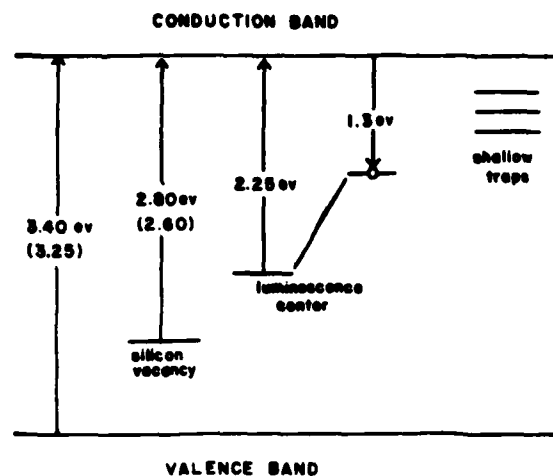


Figure 1 - Energy-band diagram of Bi<sub>12</sub>SiO<sub>20</sub> (BSO) at liquid nitrogen (room) temperatures.

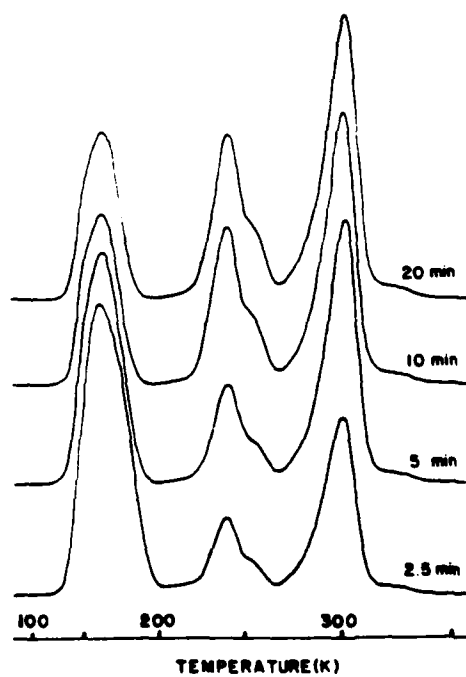


Figure 2 - Thermally stimulated current (TSC) for BSO for 2.5, 5, 10, and 20 min. irradiation. The heating rate was 0.8 K/sec.

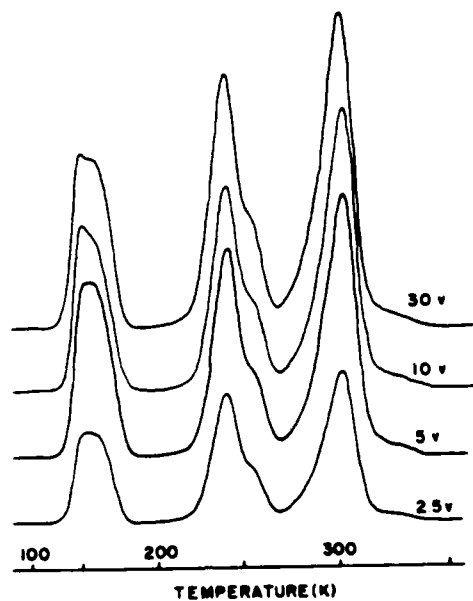


Figure 3 - TSC for BSO for 2.5, 5, 10, and 30 volt bias; the vertical scale (current) is arbitrary. The irradiation time was 10 min.; the heating rate was 0.8 K/sec.

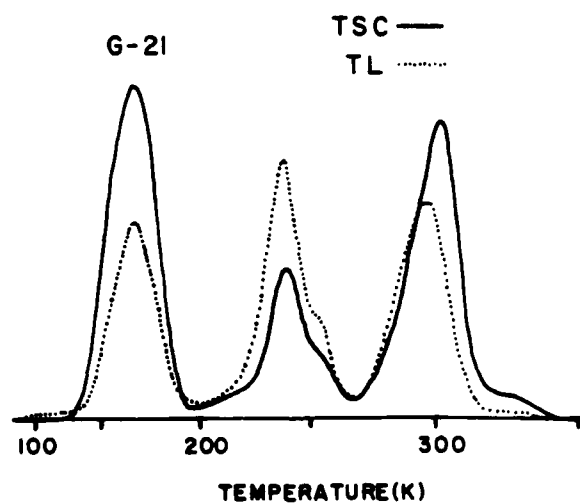


Figure 4 - TSC and TL for sample G-21.

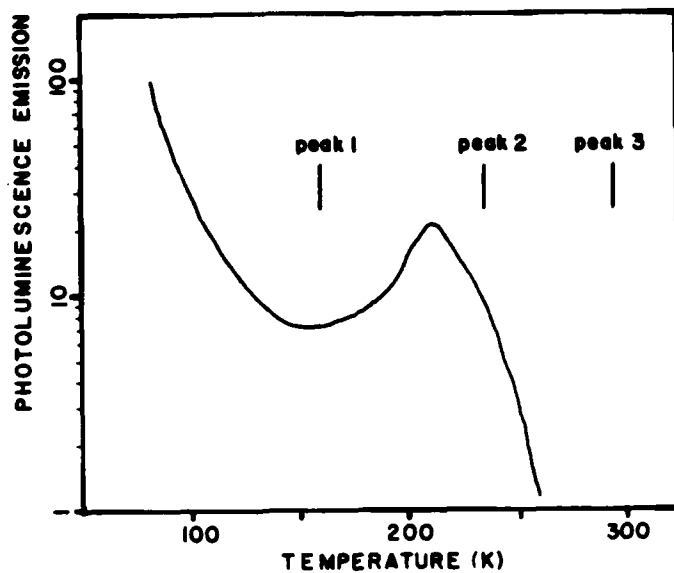


Figure 5 - Temperature dependence of photoluminescence emission at 1.30 eV in BSO under 3.4 eV irradiation. (After Hou et al.<sup>6</sup>)

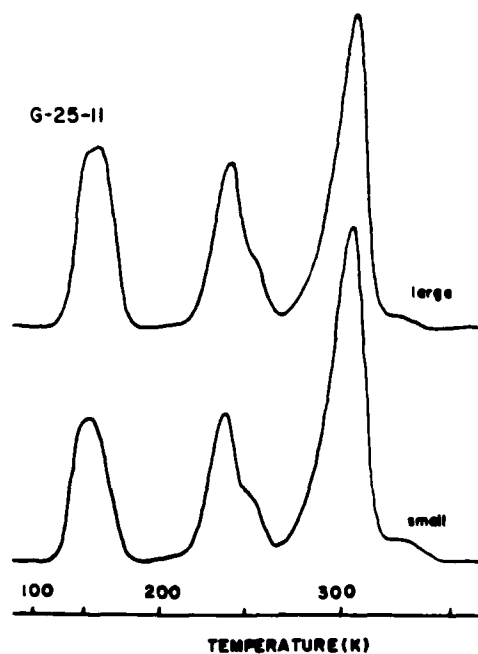


Figure 6 - TSC for small ( $0.5 \times 0.5 \text{ cm}^2$ ) and large ( $1.0 \times 1.0 \text{ cm}^2$ ) specimens of sample G-25-11.

1984 USAF-SCEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

COMPARISON OF PERIOSTEUMS FROM FEMUR AND VERTEBRAL BONE

Prepared by:	Dr. Gwendolyn B. Howze
Academic Rank:	Associate Professor
Department and University:	Department of Biology Texas Southern University
Research Location:	Aerospace Medical Research Laboratory Biodynamics & Bioengineering Division Biodynamic Effects Branch Wright-Patterson Air Force Base
USAF Research:	Dr. Leon E. Kazarian, Mr. Clarence M. Oliff
Date:	September 28, 1984
Contract No.:	F49620-82-C-0035



# Comparison of Periosteums from femur and Vertebral Bone

by

Gwendolyn B. Howze

## Abstract

Periosteum samples from lumbar and thoracic vertebrae have been studied by scanning electron microscope. Vertebral periosteum conforms to the description developed for femur periosteum during the 1983 SFRP and 1983 RJP.

Collagenous fibers are the most numerous structural compone. The collagenous fibers are composed of fibrils aligned in parallel array. The fibers are disrupted by collagenase of very high purity. In addition, a simple cryofracture technique was used to study the internal structure of the collagenous fibers and blood vessels.

A network of fine fibrils similar to that seen in femur periosteum has also been visualized. The mesh seems less plentiful in vertebral periosteum. Since treatment with very pure collagenase destroys the network, it is concluded that the network is equivalent to the reticular fibers of light microscopy. The presence of the fibrils in periosteum, and their identification as reticular connective tissue has not previously been reported in the literature.

In both types of periosteum, the osteogenic or combium layer is very cellular. A large proportion of the cells have on in situ organization which has the appearance of a stratified squamous epithelium. The fibril mesh work is usually disposed between the cell layers. Periosteum dissection by collagenase has been an effective method for studying cellular variety.

As expected, vertebral periosteum has a very rich blood supply.

#### ACKNOWLEDGEMENTS

The author would like to thank the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the opportunity to spend a profitable and interesting summer at the Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. Particular thanks is given to the Biodynamic Effects Branch (Biodynamics & Bioengineering Division) for its hospitality and quite adequate working conditions.

Among the numerous people who contributed to making this a very satisfying experience. I would like to especially thank Dr. Leon E. Kazarian and Mr. Clarence M. Oloff for suggesting an area of mutual interest which could profitably be studied in the short time period available. They also provided continued excellent support during the research period. Miss K.C. Smith also made essential contributions to the effort. Special acknowledgement is given to Dr. Gerald Taylor and Mr. Joe Dardano of the Biomedical Branch, Clinical Laboratory at the L. B. Johnson Space Center for providing very liberal access to their Scanning Electron Microscope and ancillary equipment.

## I. Introduction

This project is a continuation of research begun during the 1983 Summer Faculty Research Program (83 SFRP), in residence at the Biodynamic Effects Branch (BD) , Biodynamics and Bioengineering Division (BB) Air Force Aerospace Medical Research Laboratory (AFMRL). It was continued at Texas Southern University under the 1983 Research Initiation Program (83 RIP). It is being done in collaboration with Mr. Clarence Oloff and Dr. Leon E. Kazarian, Branch Chief, Biodynamic Effects Branch (AFAMRL/BB/BD).

The general objective remains the same, it is to determine if there are detectable structural changes in the osteogenic layer in the periosteum of bone undergoing demineralization.

The periosteum is the dense connective tissue coat which covers boney surfaces. The combium or osteogenic layer of the periosteum is the very cellular region immediately adjacent to, or appositional to the osseous matrix. The osteogenic layer contains the cellular elements which play a role in bone histiogenesis and repair of bone fractures (1, 2).

The normal or unstimulated structure of femur periosteum was studied by scanning electron microscopy (SEM) during the tenure of a 1983 Air Force Summer Faculty Research Fellowship. In two species studied, rat and Rhesus monkey, five distinct morphological regions were discerned, although sharp boundaries were not visible. The top-most fibrous layer is a coherent layer of fibers in an interknitted array. Intermixed with the fiber bundles are many blood vessels. Below the top layer is a second fibrous layer in which the fibers are more loosely arranged and mixed with a few cells. A third region contains cells, fibers and blood vessels in an

apparently random innominate distribution. Cells are plentiful in this region. The fourth layer is highly cellular. In the monkey, the cells are stratified in layers with each layer separated by a network of fibrils. The fifth region which is justa-matrix is apparently one cell layer thick and in which the cells are strongly attached to the subjacent bone matrix.

In both species, there are certain thick fibers which are composed of thinner fibrils aligned in parallel array. Because the fibrils which constitute the fibers exhibit a lengthwise structural periodicity, these fibers were tentatively called collagenous fibers. Evidence to support this conjective is presented in the results section. The fibers characteristically are covered with a meshwork, also composed of fibrils, which usually contain pariculate structures of unknown composition. Evidence is presented in this study, that all fibrils are composed of collagen.

In addition to the fibers and cells, periosteum contains a very rich vascular network.

The current study investigates the question of whether there are significant differences in the structural organization of periosteums from vertebrae and femurs. The techniques of cryofracture and enzyme mediated tissue dissociation have been very useful to this study.

The studies, which have been completed to date, have attempted to compile a description of periosteum which can serve as a baseline for the projected demineralization studies.

#### The Rationale

The rationale for this research consists of three simple hypothesis.

- 1) During the process of bone demineralization, there are coincident and

and related events occurring in the periosteum and endosteum. 2) The putative demineralization related events are manifested as structural changes in the constituent elements such as cells, fibers and blood vessels. 3) The anticipated periosteal changes can be visualized using the scanning electron microscope.

#### Desired Outcome

It is hoped that the collected observations and results will lead to testable predictions about the mechanism of bone demineralization.

#### Relevance

This project has relevance to bone demineralization which occurs during space flight, conditions of weightlessness or low gravitational force, and conditions of prolonged immobility.

## II. Objective

The overall objective of this project is to investigate the possibility that the osteogenic region, or combium, of the periosteum plays a significant role in the bone demineralization which is induced by hypokinesia or immobilization.

This goal is being approached in two stages. Stage I was carried out in the summer of 1983 at the Air Force Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base. The stage I studies consisted of a scanning electron microscopical (SEM) study of normal periosteum of long bones, such as femur and tibia, from both rat and Rhesus monkey. The specific aim was to obtain a baseline description as prerequisite to any attempt to detect experimentally induced changes in periosteal structure.

The obligate prerequisite for our study is a detailed knowledge of normal periosteal structure. The research which has been done, to date, has been to generate the required baseline description of periosteum which was not available in the technical literature. We have generated a collection of data describing periosteum of untreated long bones such as femur (thigh) and tibia (shin). Recently, however, the group at BD has redefined the project in terms of the lumbar vertebrae. Engineering principles dictated this modification, the vertebral column is the major weight bearing structure in humans. This shift of emphasis to the vertebral bones and the selection of the lumbar vertebrae as the model bones for studying bone demineralization has certain consequences. In the first place, it is a more valid approach from the standpoint of bioengineering. Secondly,

it requires an answer to the question of whether there is a significant difference in the organization of periosteum from long bones (which we have been studying) and periosteum from lumbar vertebrae, the current bones of interest. Consequently, the baseline study was extended to include periosteum from lumbar vertebrae. A second reason for extending the baseline study was the introduction of a new and improved biopsy procedure for obtaining periosteal specimens. It needs to be shown that the new method does not introduce artifacts. The specific aims of the research done during the 1984 SFRP are described below.

#### Specific Aims

The primary aim is to determine if the structure of vertebral periosteum is significantly different from long bone (femur and tibia) periosteum. Additionally, it was necessary to acquire more information on the types of cell population, and to determine the chemical composition of the fibers.

The secondary aim is to determine if the biopsy sampling procedure induces any structural anomalies. The questions which must be answered are the following. 1) Does the cutting tool (used in the biopsy) cause changes in the periosteum surrounding the entrance and exit sites? 2) Does the entrance site differ from the exit site? 3) Does regeneration occur at the sampling sites.

### III. Research Plan

#### The Approach

The types of experiments which were done to achieve the specific aim are described in this section.

1) Lumbar vertebral periosteum specimens were taken from Rhesus monkeys using a biopsy technique developed by Mr. Clarence M. Oloff, BEB/BBD/AFAMRL.

2) Scanning Electron Microscopy (SEM)

2.1) The specimens were chemically processed according to methods developed in the 83 SFRP and the 83 RIP and described in the summary of methods section below.

2.2) The chemically processed specimens were studied by SEM.

3) The in situ distribution of cells and fibers in the internal regions of the periosteum: While it is very easy to study the two surfaces of periosteum; study of the regions between the two surfaces, may result in the production of misleading anomalies. One method which is thought to be less artifactual than most, is that of freeze fracture. In this method, after freezing the specimen, one induces a fracture along the most susceptible interface.

3.1) Freeze fracture: vertebral periosteum was fixed as described in the summary of methods section, it was subsequently frozen and fractured. The fractured surfaces were coated with gold and studied by SEM.

4) Cellular populations in the periosteum

4.1) Isolation and harvest of cellular types: unfixed periosteum



was dissociated with crude collagenase and the cells were harvested by filtration onto millipore filters. The sedimented cells were fixed. Samples were smered and stained for light microscopy (LM), and samples were processed for SEM.

- 4.2) Microscopical study of harvested cells to determine cell types and relative proportions of the types. SEM and LM were used.
- 5) Chemical characterization of fibers by enzyme hydrolysis (Graduate student)

Enzyme specificity was used to detect and identify certain molecules. The rationale being: if certain molecules are integral to the continuity of a structure, hydrolysis of a sufficient quantity of the molecules will undermine or disrupt the structure in question. To be specific, if very pure collagenase causes the disruption of certain periosteal fibers, then it can reasonable be assumed that the major constituent of the fiber is collagen.

- 5.1) Unfixed periosteum was treated with collagenase of very high purity for specified incubation periods. The reaction was then stopped. The treated periosteum was fixed and processed for SEM. The processed samples were studied by SEM.

### Summary of Methods

In order to achieve the stated objective, standard scanning electron microscopical procedures (4, 8) have been slightly modified to suit the periosteum. The sequence of steps in summary are the following:

- 1) obtaining the specimen. 2) preservation of native state by fixation.
- 3) dehydration. 4) preparation of SEM. 5) observations and producing a permanent record

Obtaining the specimen: The rats were sacrificed and the femurs were dissected out, stripped of muscles and washed in buffered saline. The Rhesus monkey periosteum were obtained by biopsy. In spite of a pre-fixation wash, the Rhesus samples were contaminated by adhering red blood cells.

Fixation: The specimens were fixed in 2% glutaraldehyde in pH 7.3, 0.2 M phosphate from 3 to 24 hours.

Dehydration: A graded series of ethyl alcohol was used in the dehydration, 35%, 70%, 95%, and 100%.

Preparation for SEM: Normally after the dehydration, specimens are transferred to iso-amyl acetate in preparation for the critical point drying procedure. Carbon dioxide was used for the critical point drying. Next the specimens were attached to the stubs with a conductive cement and covered with a gold coat estimated to be 20nm thick. The observations were made on an ETEC Autoscan Scanning Electron Microscope. Uniformly, the accelerating voltage was 20 Kv. A permanent record was obtained upon Polaroid 55 P/N film. This procedure was modified slightly for rat specimens.

Modified procedure for rat femur: This procedure was developed for the

rat femur, but it probably should be used uniformly since it eliminates certain artifacts, e.g. dissection injury and shrinkage artifacts due to fixation and dehydration. The sequence is as follows: 1) the whole femur is fixed. 2) the dehydration was the same as above, but a sub-routine to separate the periosteum from the matrix was inserted before the critical point drying procedure. 3) isolation of the periosteum by embedding it in a collodian membrane. The details of this procedure are listed below. 4) critical point drying and continuation with the standard SEM procedure.

1. The bones were placed in a 1:1 mixture of absolute ethanol and amyl-acetate, two changes, 15 min. each.
2. Next, the bones were placed in 1.5% collodion (dissolved in the ethanol/amyl acetate solvent) for 15 minutes.
3. They were next transferred to a small container of 5% collodion for 5 minutes.
4. The top of the containers were then removed as the solvent evaporated for 5 min.
5. The bones were removed and held in the air for a minute or more until a definite solid film could be seen, and the smell of solvent had decreased. Caution: do not allow it to dry completely, rotate to insure a uniform film.
6. The semi-dry sample was placed in a container of distilled water for 5 min.
7. A circular cut (through the film down to the ma-

trix) was made at each end. One longitudinal cut, extending between the two circular cuts, was made.

8. The bone was split along the longitudinal line.
9. The membrane was peeled off the split bone, either as one or two pieces.

Comments: The methods described above were developed during the tenure of the 1983 SFRP. The following section consist of additional modifications and new methods.

New Methods: The freeze fracture procedure (18)

1. Fix the specimen as described above.
2. Process the specimen through ethyl alcohol graded series.
3. After the final 100% ethyl alcohol treatment, insert specimen into small cylinder of parafilm contains 100% Ethyl Alcohol.
4. Crimp shut both ends of the cylinder.
5. Hold cylinder under liquid nitrogen until frozen.
6. Place the frozen cylinder containing the frozen tissue on a metal block precooled in liquid nitrogen.
7. Fracture the cylinder with a precooled single edge razor blade.
8. Thaw the specimen in fresh absolute ethanol, and remove from the parafilm cylinder.
9. Dry sample by critical point procedure using 100% ethanol, rather than amyl acetate.
10. The dried tissue fragments should be guided onto specimen stubs with silver conductive cement. The fractured surfaces (which are distinctly smooth and shiner when viewed under a dissecting microscope) are oriented upward.

11. Coat specimen on a rotating stage by vaporizing gold to deposit a layer ca. 20nm thick.

The following two methods employ the enzyme collagenase, in one instance to release cells, in the second case to detect the chemical presence of collagen in fibers.

New Method: The cell isolation procedure (19, 20)

1. Remove femur as described above.
2. Score diaphysis in several spots.
3. Wash bone in Gey's salt solution.<sup>a</sup>
4. Place washed bone into 3cc of dissociation fluid.<sup>b</sup>
5. Incubate in water bath at 37 degrees C for 30 min.
6. Vortex at intervals of 15 min. for total of 1 hr. If cells are present in sufficient amount at end of 45 min. decant cell suspension and chill.
7. Add 2ml of NaEDTA (0.1M, pH 7.4) to the first fraction, and continue to chill.
8. Put bone in 3cc of fresh dissociation fluid and continue to incubate vortexing every 15 min., do not exceed 1 hr.
9. Add second fraction to the first fraction after a sufficient number of cells accumulated.
10. Centrifuge at lowest position<sup>c</sup> for 5 minutes or until a soft pellet is formed. Decant the supernant.

---

<sup>a</sup>Gey's balanced salt solution was purchased from GIBCO Laboratories.

<sup>b</sup>Gey's balanced salt solution containing 2mg/ml of collagenase (Sigma type IV), Glucuronidase (Sigma) 1mg/ml hyaluronidase (Sigma) and 1mg/ml trypsin (Sigma).

<sup>c</sup>IEC Clinical centrifuge, desk top model.

11. Vortex the pellet gently.
12. Add 1ml of Gey's salt solution, vortex gently until cells are suspended.
13. Collect on millipore filters or make smears on microscope slides.
14. Fix the cells on the millipore filters with 2% Glutaraldehyde in 0.2M, pH 7.3 phosphate. Fix the smears with 3:1 absolute alcohol and glacial acetic alcohol.
15. Dehydrate the preparation.
16. Process the microscopy.

New Methods: Chemical characterization of fibers by enzyme treatment.

1. To six ml. of enzyme solution<sup>a</sup>, add small pieces of fresh monkey periosteum or fragments of fresh femur.
2. Incubate at 37 degrees C.
3. Fix samples at intervals of 30 min., one hour, two hours.
4. Dehydrate specimens.
5. Process for SEM: use the method described above as "Modified procedure for rat femur".

---

<sup>a</sup>600 units of collagenase (Sigma Type VII, EC.3.4.24.3) per 6 ml of Gey's balanced salt solution.

#### IV. Results

The external surface of periosteum consists of a coherent interlocking array of fibers. Figure one is from rat femur. Rhesus femur periosteum is almost identical, in the micrograph available, red blood cell contamination obscures the detail. Figure one and two essentially identical in organization. Figure two is of Rhesus vertebral periosteum. The fibers are actually bundles of fibrils aligned in parallel array. The thickness of the subunit fibrils range from 60 nm to 100 nm (the thickness of the gold coat was subtracted).

Figures three and four illustrate the fibrillar organization of the fibers. Figure three gives a surface view. Figures four, five, and six provide a cross-sectional view due to cryofracture.

The fiber bundles are usually enclosed in a reticulum of fibrils which I have referred to as the fibril network. The fibril network is illustrated in figures 7 and 8, vertebral and femur respectively.

Figures 9A and 9B show internal structure of blood vessels as revealed by the cryofracture technique, they are probably arteriole and small vein respectively. An intact artery is shown in figure 10.

Treatment of fresh rat femur periosteum with Sigma collagenase type IV (high purity) disrupts the fibrils of all types. After one hour all of the fibrils have been digested leaving only cell sheets intact. Fig. 11 is a typical scene which is seen after prolonged exposure to dilute collagenase solution. It would appear, therefore, that all of the fibrils contain collagen. The enzyme specificity and the banding pattern support this conclusion. Since the fibrils of the network

are also sensitive to collagenase, the current conjecture is that they are the SEM equivalent to the reticular fibers of light microscopy.

Some of the cell types seen in femur periosteum are shown in figures 12 through 19. Figure 12 shows a typical cell type associated with collagen fibers. The cells at the arrow are magnified in figure 13. Cells with very visible microvilli are shown in figures 14 and 15. The cell indicated by the arrow is magnified in figure 15. An area containing layered sheets of flattened cells is shown in figure 16. This organization is reminiscent of stratified squamous epithelium. Note the presence of the fibril network between layers of cells. Cells with long processes are shown in figures 17 and 18. The cells in figure 19 also exhibit processes. They are also covered or enclosed by a material which produces a hardened or stiff appearance. This scene is interpreted to be a site of lamella bone formation. The cells shown in the following micrographs were found in Rhesus vertebral periosteum.

Figure 20 shows a scene very similar to that in figure 19. Figure 20 is from Rhesus vertebral periosteum.

The center cell in figure 21 is a cell with highly developed processes. The processes at the arrow are folded under the cell. The cell is additionally joined by peg-like junctions to the cell on the right. Figures 21, 22, 23 show areas of densely packed cells with highly developed packed and anastomosing processes.

Figure 24 shows a fractured region. The cells which are seen in the right half of the field were formerly covered by a layer of fibers a fragment of which is still visible on the left.



The cellular region is magnified in figure 25. Figures 26 and 27 are the vertebral equivalent of figure 18. They show the same layered cellularity. Figure 27 also shows the distribution of the fibril network among the cell layers. The cells in figure 28 are probably the vertebral counterpart of the cells seen in figures 14 and 15. The cell type in figure 29 is usually found in association with the collagenous fibers, it is similar in shape to the cell type seen in figure 13. Figures 30 and 31 are examples of other cell types found in situ in vertebral periosteum.

Figures 32 to 41 are micrographs of cells which were released by enzymatic dissociation. Their most striking feature is the scarcity of cellular processes. Except for the cells in figure 35, 36, 37, processes are absent. Some of the cells exhibit surface features which suggest that the processes were amputated during the isolation procedure, possibly by the non-specific protease activity found in crude collagenase preparations. The cells in figure 32 and 33 are very large. They are each, at least 10 times larger than the cell in figure 35. It is not known if these cells are of the same type. Such large cell, however, have been visualized by light microscopy in fixed and stained smears. The light microscopical studies will be described in a separate manuscript, presented by C. Ruben, the graduate student participant.

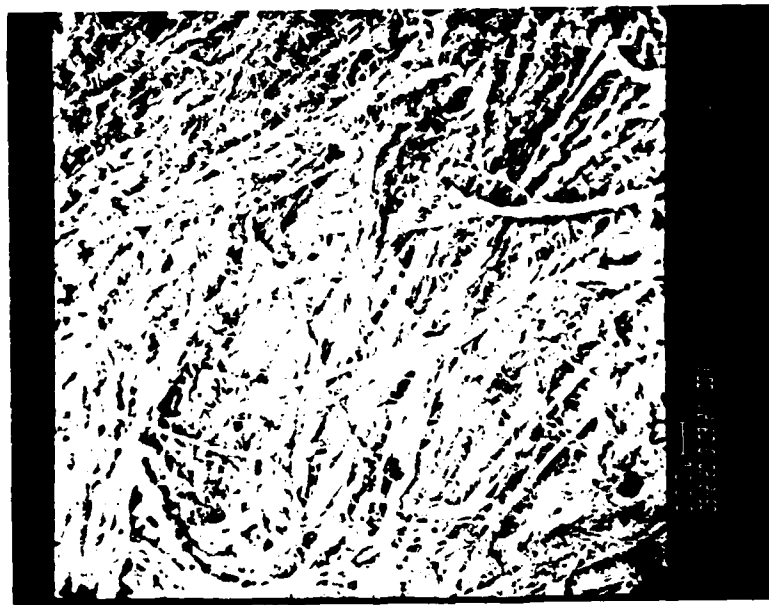


FIGURE 1 - EXTERNAL SURFACE  
RAT FEMUR PERIOSTEUM

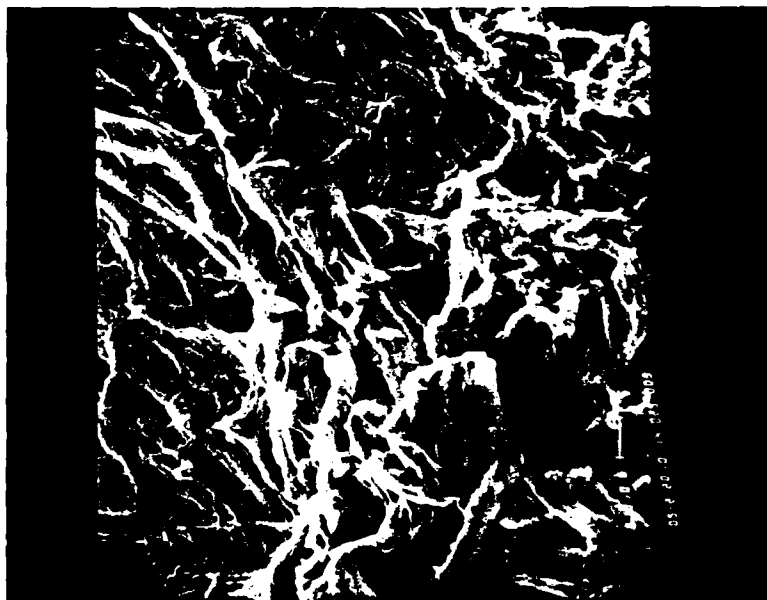


FIGURE 2 - EXTERNAL SURFACE  
RHEBUS VERTEBRAL PERIOSTEUM

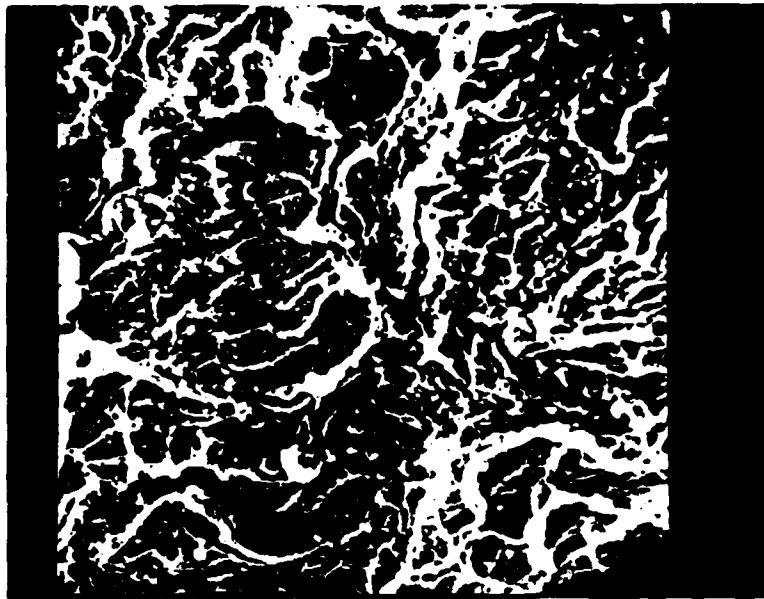


FIGURE 3 - SUBUNIT FIBRILS  
RHESUS FEMUR PERIOSTEUM  
SURFACE VIEW

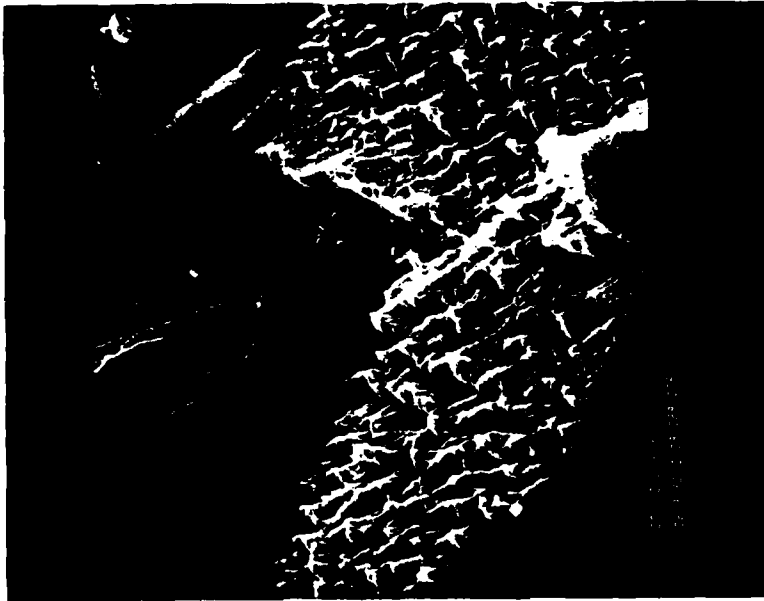


FIGURE 4 - SUBUNIT FIBRILS  
RHESUS VERTEBRAL PERIOSTEUM  
CRYOFRACTURE, CROSS SECTION

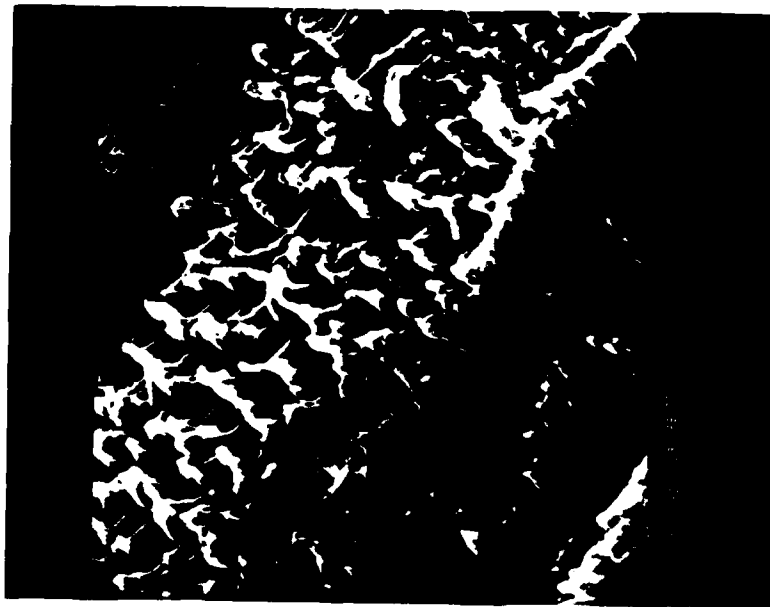


FIGURE 5 - SUBUNIT FIBRILS  
RHESUS VERTEBRAL PERIOSTEUM  
CRYOFRACTURE, CROSS SECTION

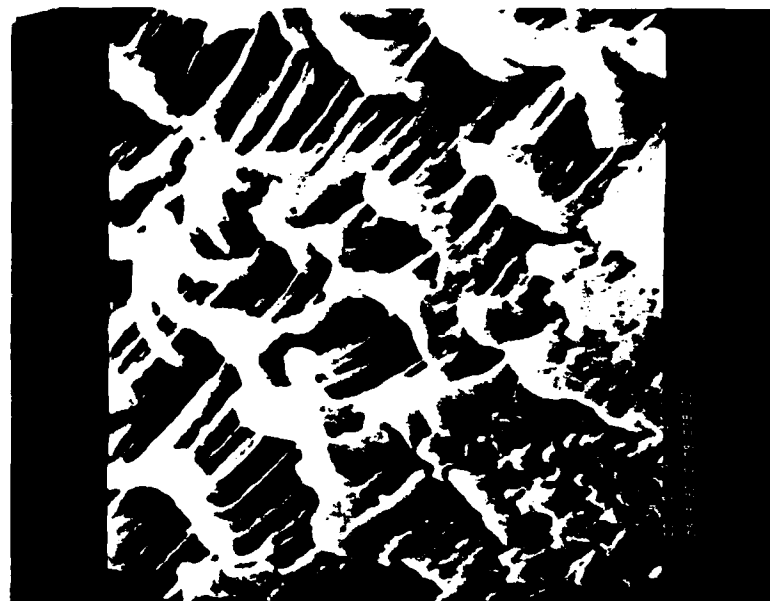


FIGURE 6 - SUBUNIT FIBRILS  
RHESUS VERTEBRAL PERIOSTEUM  
CRYOFRACTURE, CROSS SECTION

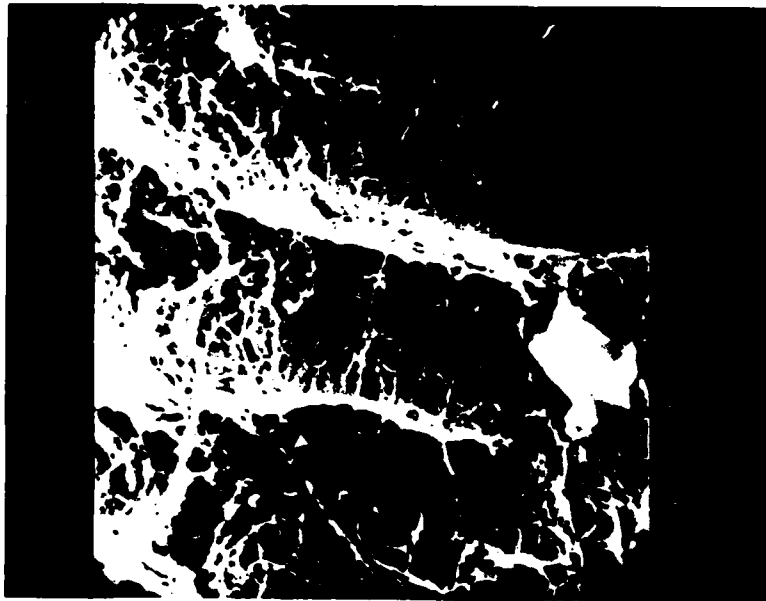


FIGURE 7 - FIBRIL NETWORK  
RHEBUS VERTEBRAL  
PERIOSTEUM

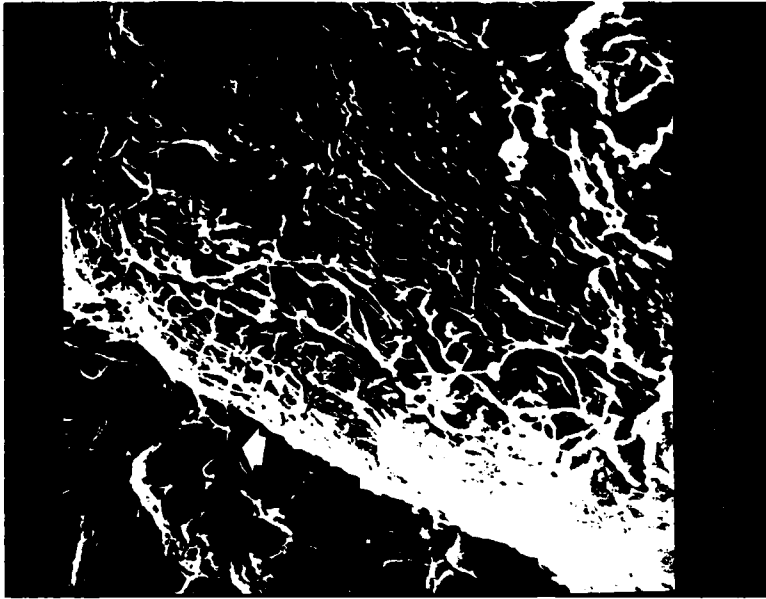


FIGURE 8 - FIBRIL NETWORK  
FEMUR PERIOSTEUM

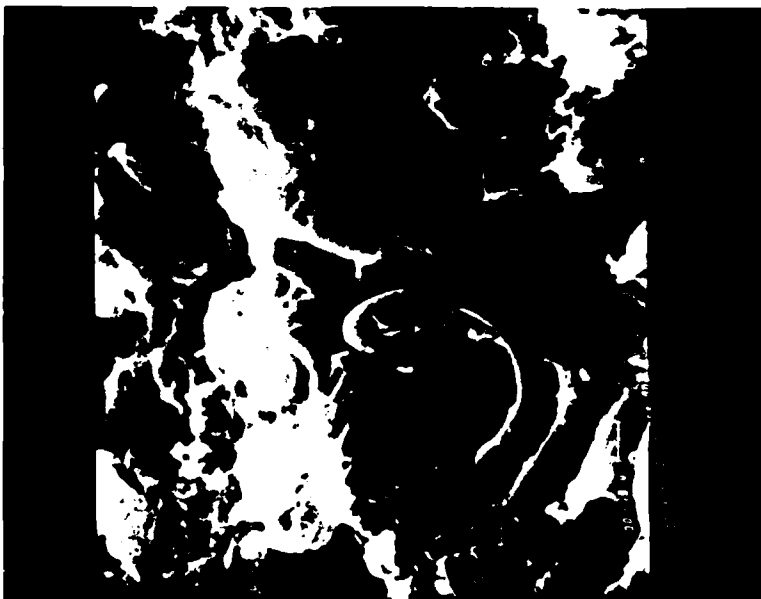


FIGURE 9A - SMALL BLOOD VESSEL  
INTERIOR  
RED BLOOD CELL at arrow  
CRYOFRACTURE



FIGURE 9B - SMALL VEIN  
INTERIOR  
LARGE CELL (arrowhead)  
CRYOFRACTURE

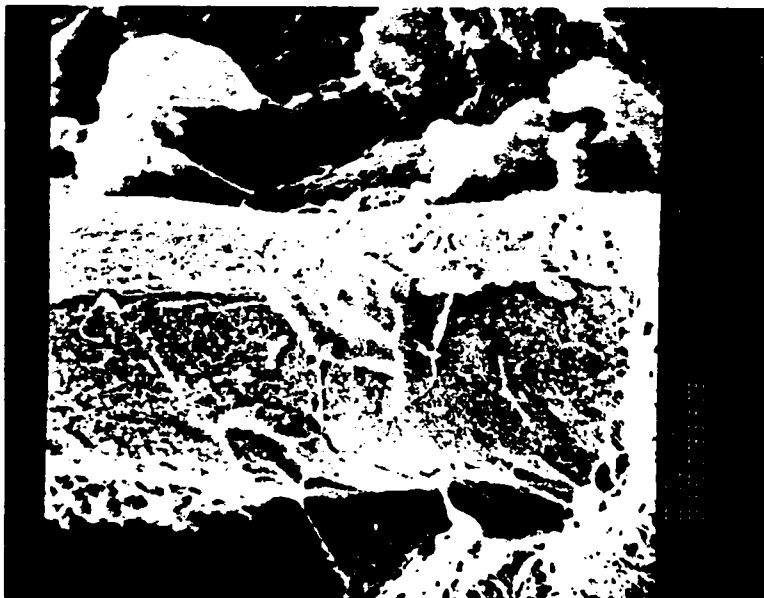


FIGURE 10 - ARTERY ENTERING PERIOSTEUM  
(dot indicates artery)  
COLLAGENOUS FIBERS RUN  
ALONG ARTERY

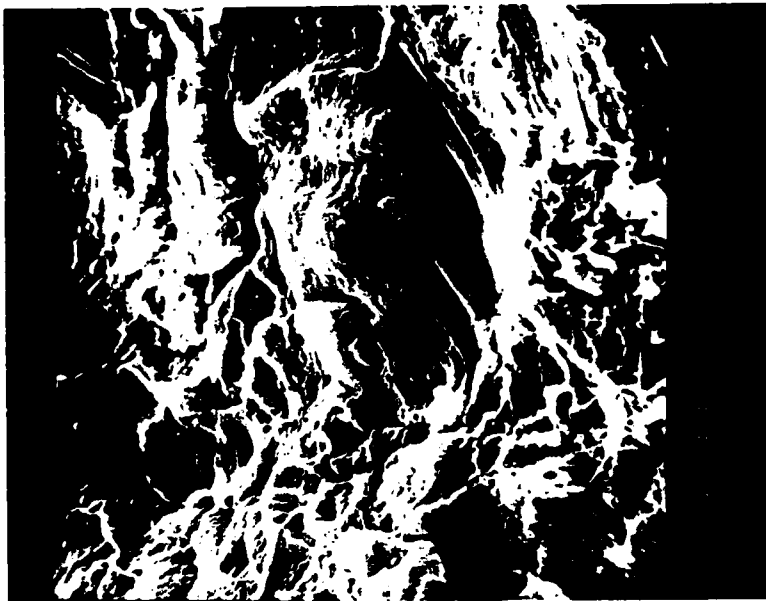


FIGURE 11 - LOSS OF FIBERS DUE TO  
COLLAGENASE DIGESTION

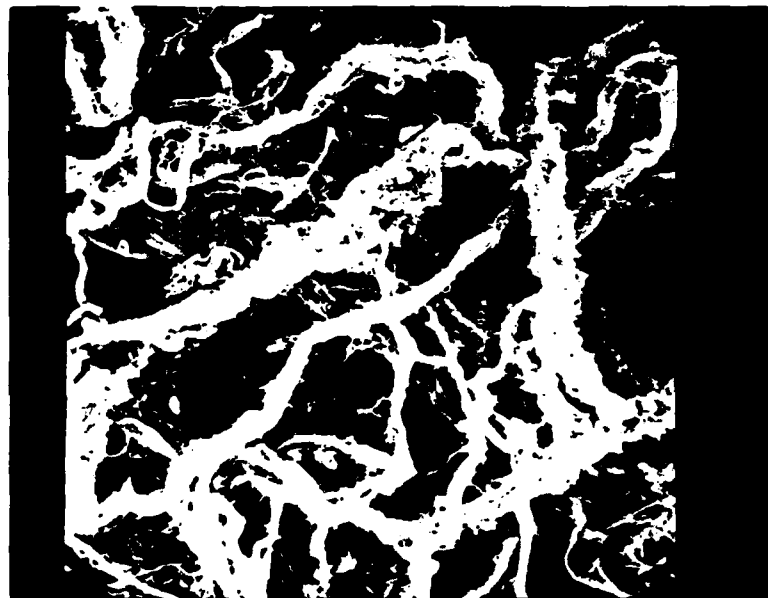


FIGURE 12 - CELLS ATTACHED TO FIBERS  
(arrow)

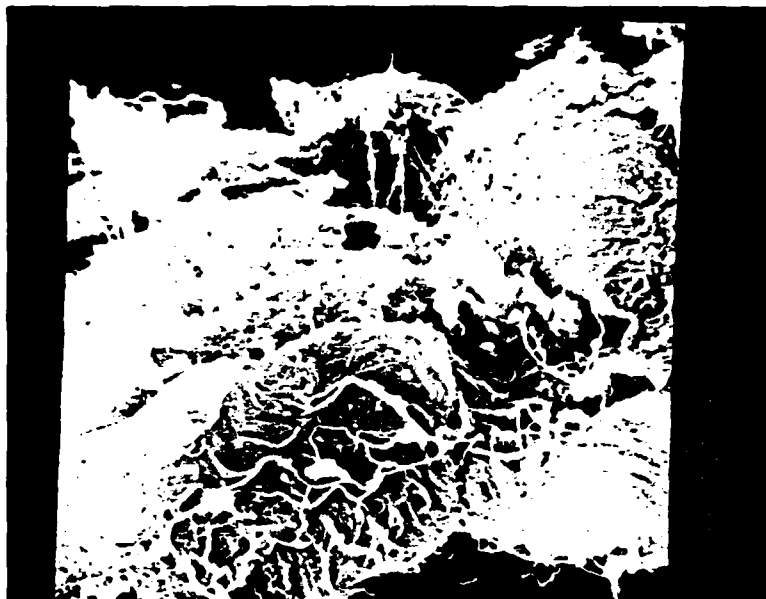


FIGURE 13 - MAGNIFICATION OF CELL IN  
IN FIGURE 12 (arrow)



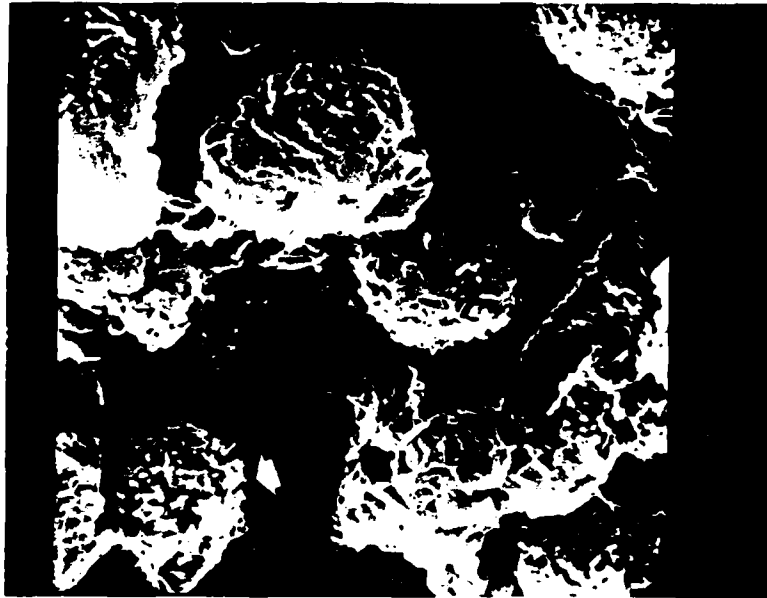


FIGURE 14 - CELLS WITH MICROVILLI

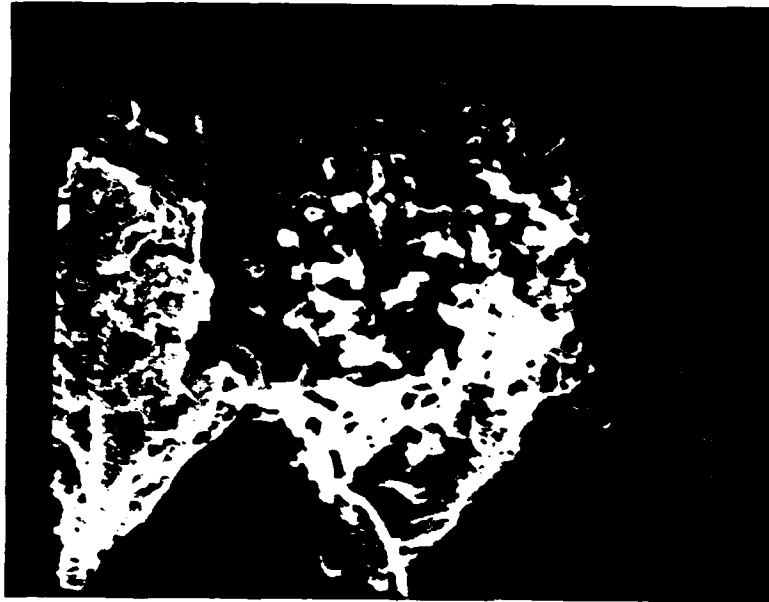


FIGURE 15 - DETAIL OF ONE CELL  
(figure 13 at the arrow)

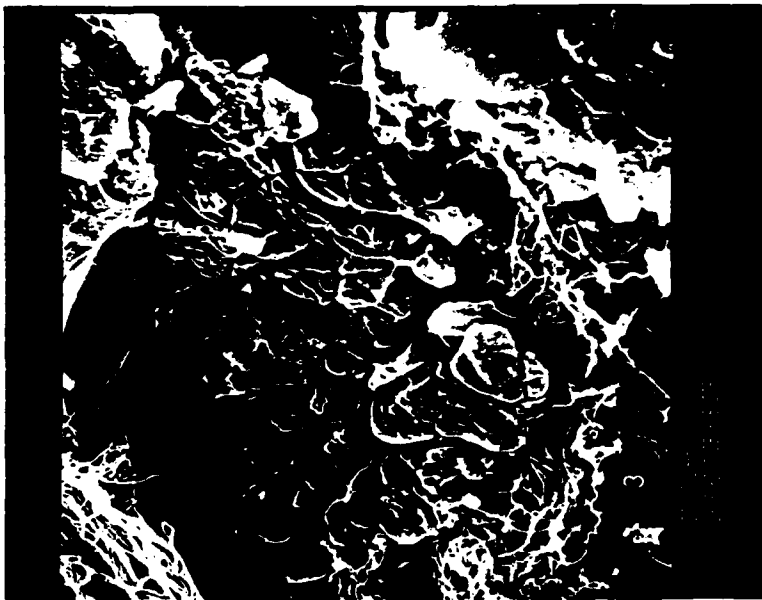


FIGURE 16 - THREE LAYERS OF CELLS  
FIBRIL NETWORK (arrow)



FIGURE 17 - CELLS WITH PROCESSES  
(arrow)

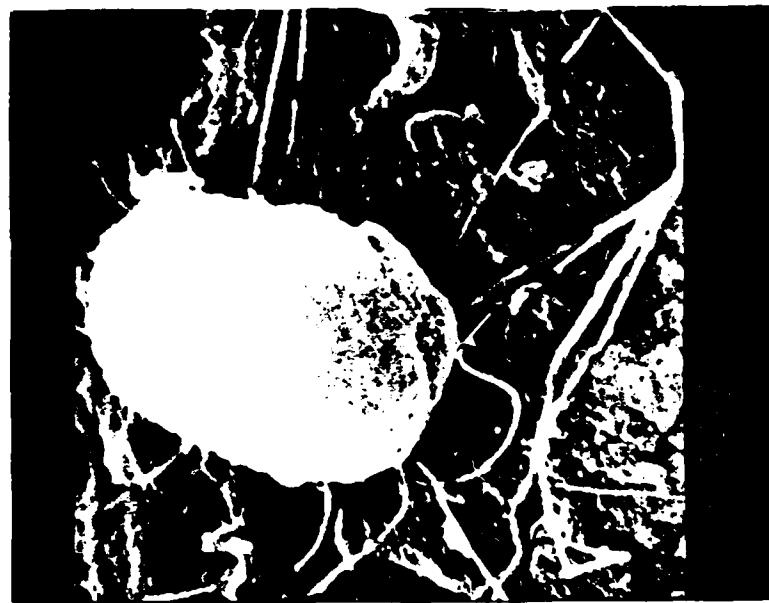


FIGURE 18 - A CELL WITH PROCESSES  
(arrow)



FIGURE 19 - BONE FORMATION (??)  
CELL PROCESS (arrow)

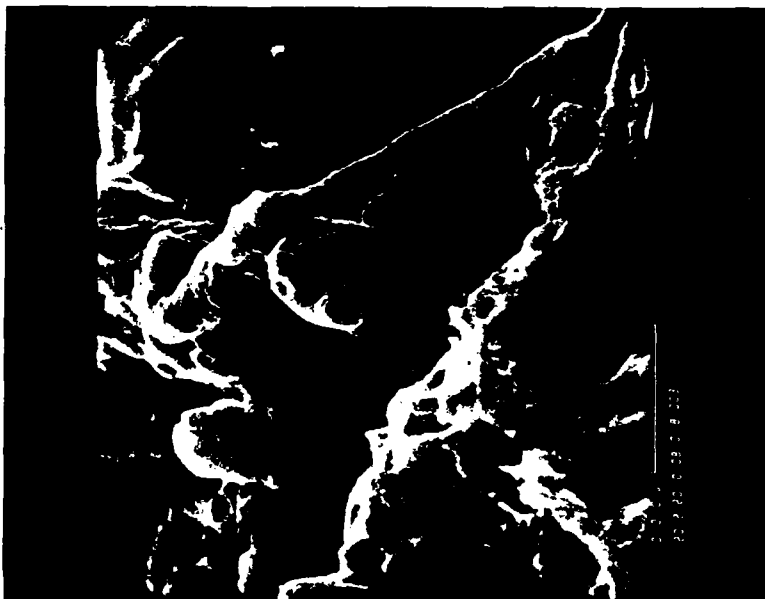


FIGURE 20 - THE VERTEBRAL EQUIVALENT  
OF FIGURE 19



FIGURE 21 - PACKED CELLS WITH LONG  
FOLDED PROCESSES (arrow)



FIGURE 22 - ANASTOMOISING AND FOLDED  
CELL PROCESSES PRODUCE  
ODD PATTERNS

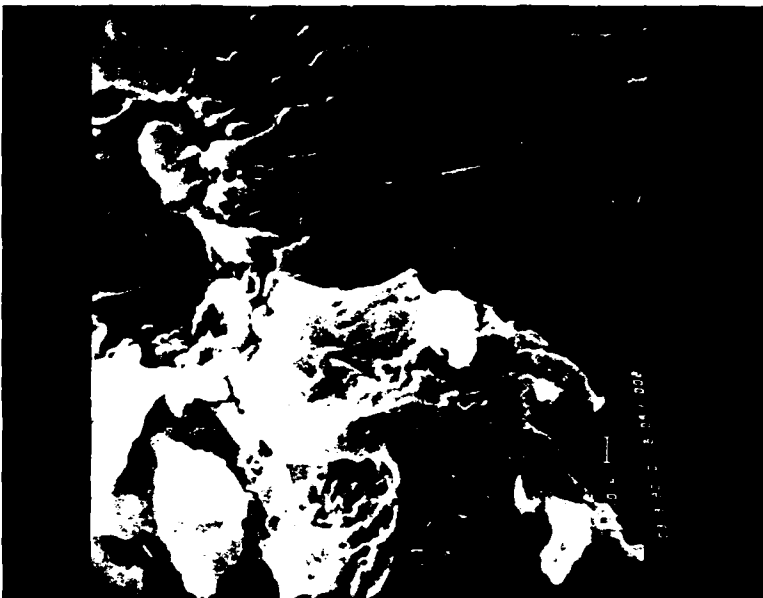


FIGURE 23 - ENLARGEMENT OF AREA AT THE  
POINTER

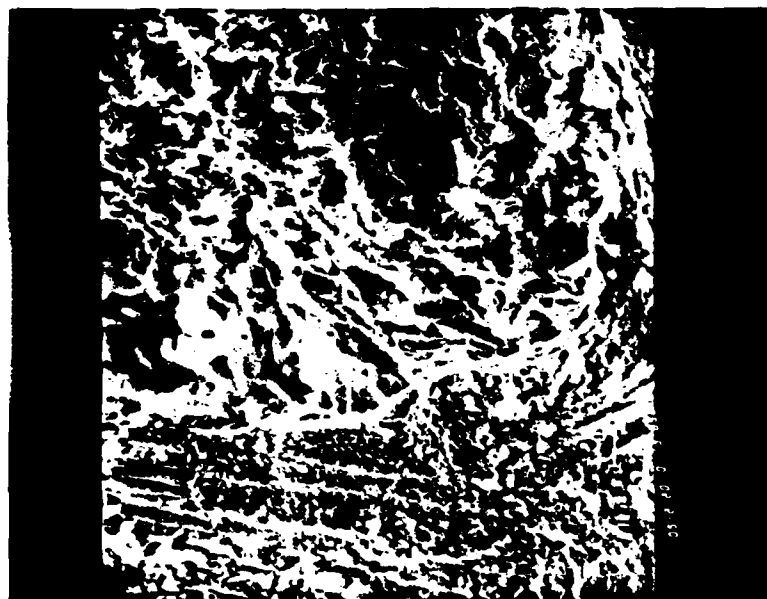


FIGURE 24 - CRYOFRACTURE  
FRACTURE OF FIBROUS LAYER  
(left) AND SUBJACENT CELLULAR  
LAYERS (right)

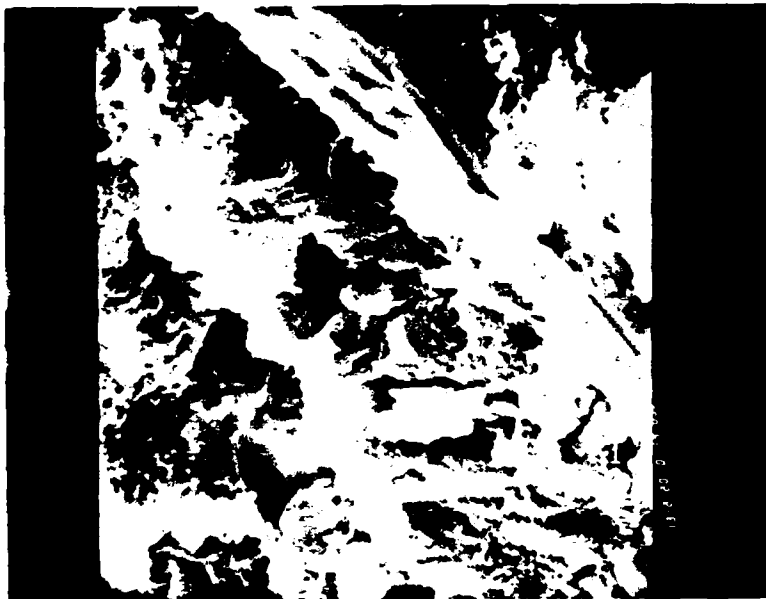


FIGURE 25 - ENLARGEMENT OF CELLULAR  
REGION

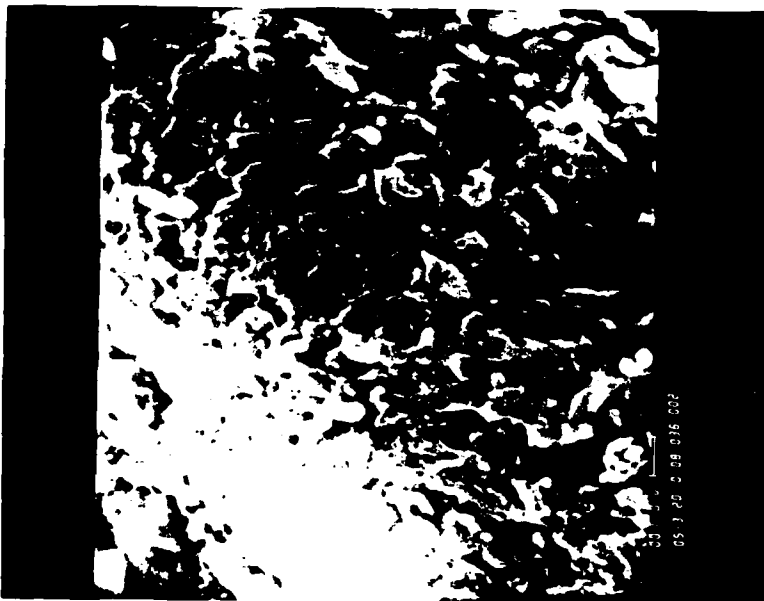


FIGURE 26 - THE VERTEBRAL EQUIVALENT  
OF FIGURE 18

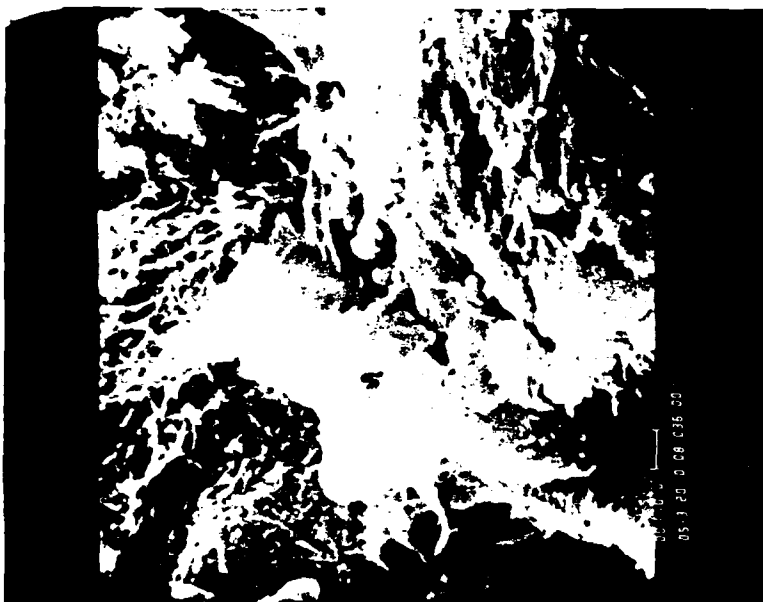


FIGURE 27 - SIMILAR TO FIGURE 26



FIGURE 28 - THE VERTEBRAL EQUIVALENT OF  
FIGURES 14 and 15

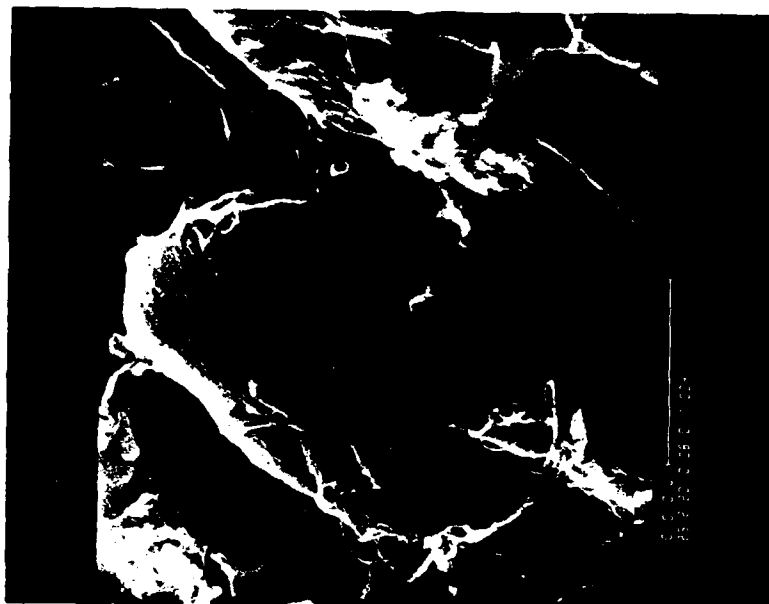


FIGURE 29 - A CELL TYPE FOUND IN  
ASSOCIATION WITH FIBERS





FIGURE 30 - CELL AND FIBRIL NETWORK  
THE POINTER INDICATES AN  
ERYTHROCYTE



FIGURE 31 - ANOTHER CELL TYPE

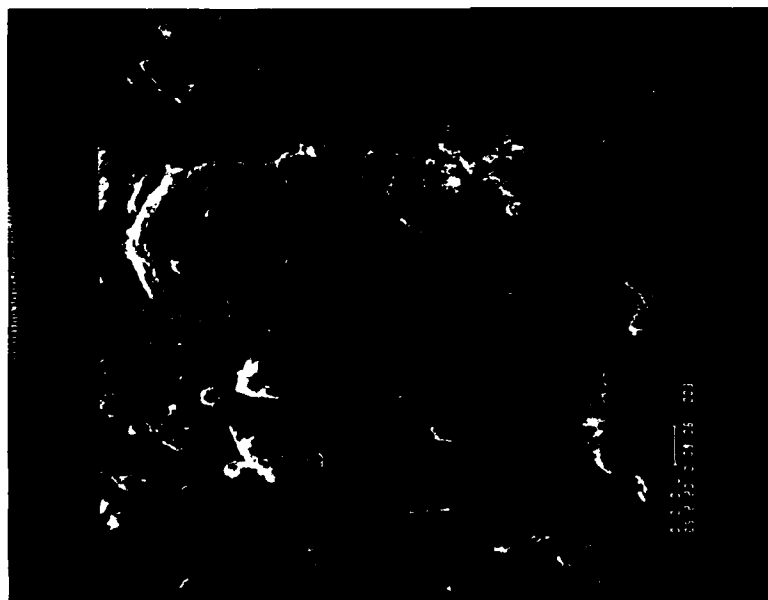


FIGURE 32 - VERY LARGE CELL TYPE  
ENZYME DIGESTION EXPERIMENT

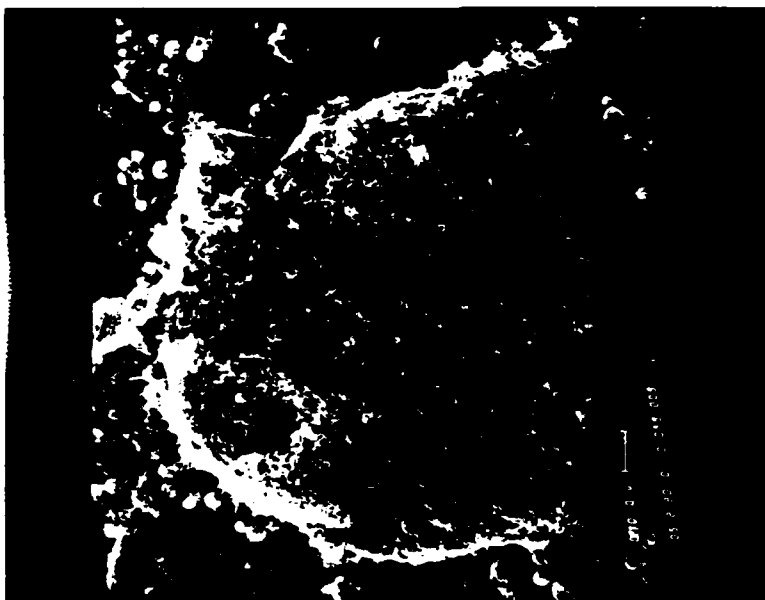


FIGURE 33 - CELL SIMILAR TO FIGURE 32



FIGURE 34 - A DIFFERENT CELL TYPE RELEASED BY  
ENZYME DIGESTION

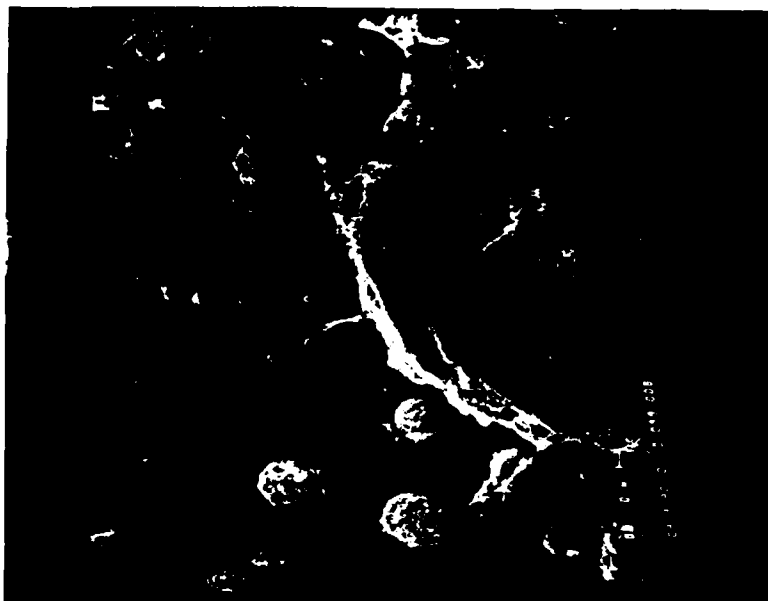


FIGURE 35 - TWO CELLS RELEASED BY  
ENZYME DIGESTION AND  
BLOOD CELLS

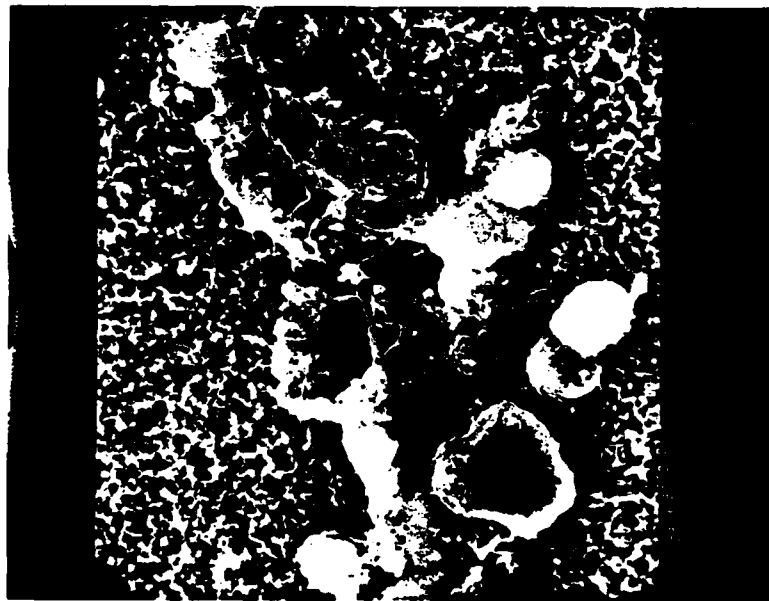


FIGURE 36 - SEVERAL TYPES OF CELLS  
RELEASED BY ENZYME ACTION

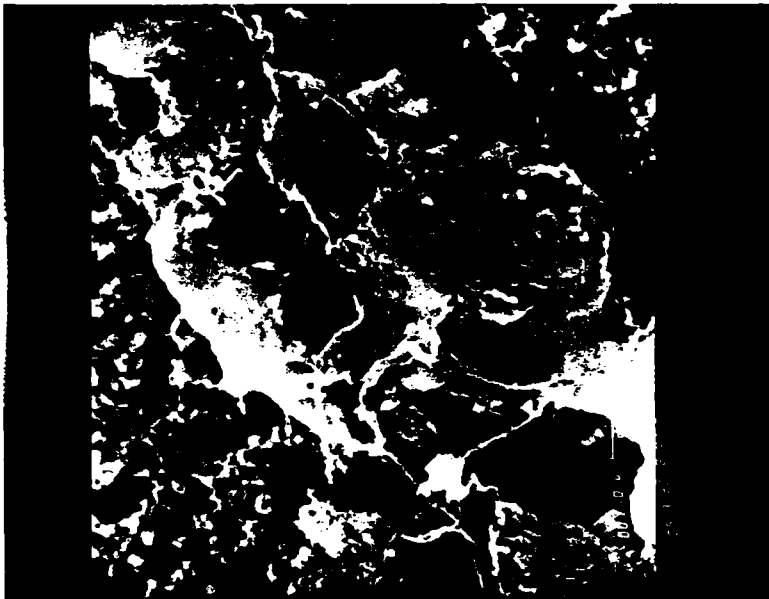


FIGURE 37 - CELL WITH PROCESSES  
ENLARGEMENT OF FIGURE 36

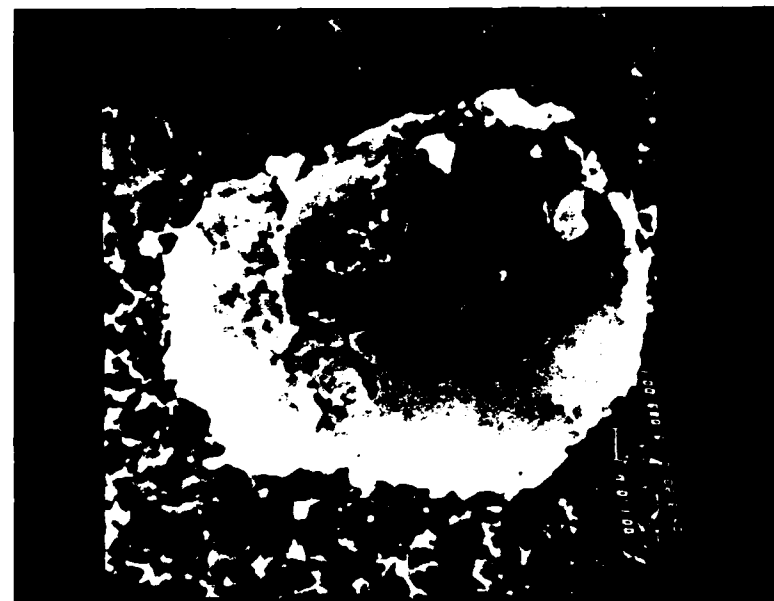


FIGURE 38 - CELL WITH AMPUTATED PROCESSES

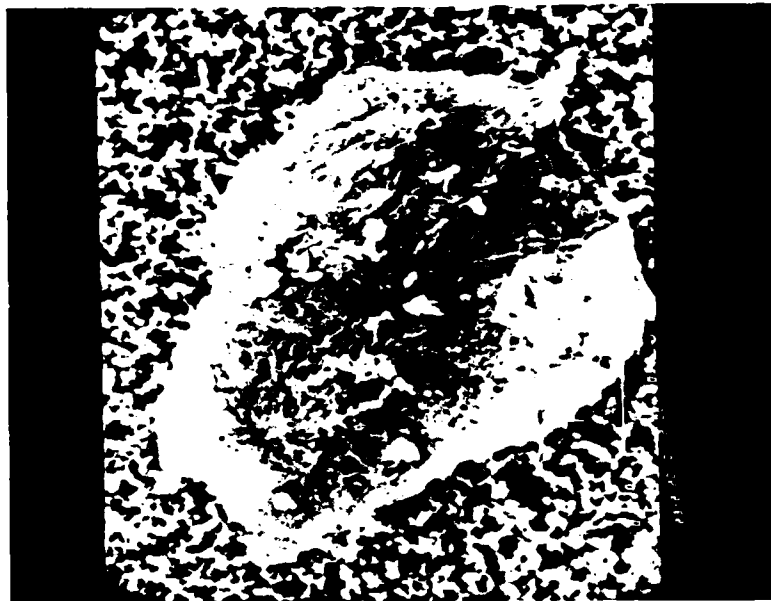


FIGURE 39 - CELL WITH AMPUTATED PROCESSES

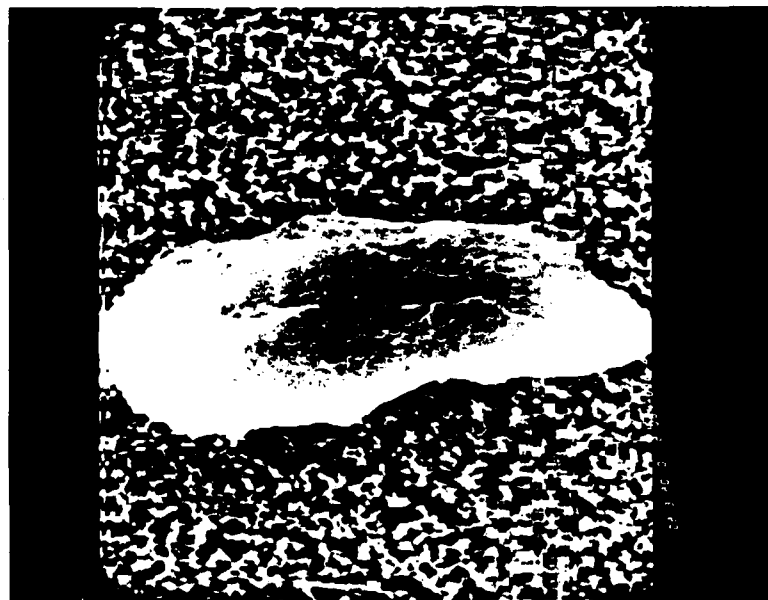


FIGURE 40 - A CELL TYPE RELEASED BY  
ENZYME ACTION

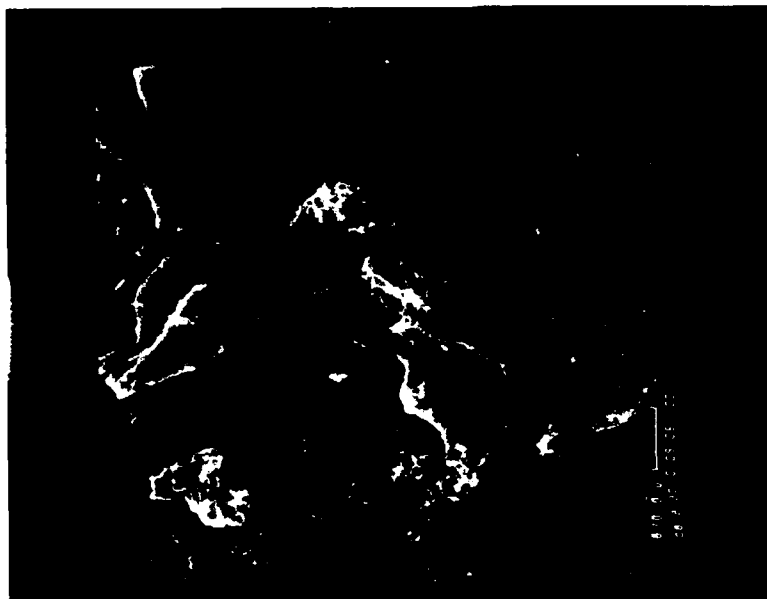


FIGURE 41 - A CELL RELEASED BY  
ENZYME ACTION

## V. Recommendations

The observations presented in the results section of this report will be used in conjunction with the published literature to produce a detailed description of the structural organization of the periosteum.

The proposed detailed description will serve as a baseline for studies of changes in periosteal structure coincident with various real life stress and/or experimental conditions which cause change in the bone matrix. In particular, it can serve as a guide as to which components are likely candidates for close observation when looking for periosteal changes coincident with bone matrix demineralization induced by hypokinesia; or matrix changes due to vibrations or impact. On the basis of the present study: 1) the possibility of changes in the number of and structure of the cytoplasmic processes on the cells in the "osteogenic" layer should be considered; 2) the ubiquitous fibril network is a likely specimen for study; 3) changes in the cell population associated with blood vessels should also be studied.

The Biodynamic Effects Branch of AFAMRL/BBD is interested in the three "bone-stress-activities" listed above.

#### REFERENCES

1. Ham, A.W. and D. H. Cormack, Histophysiology of Cartilage, Bone, and Joints, Physiology, J. B. Lippincott Company, 1979
2. Ham, A. W. and W. R. Harris, "Repair and Transplantation of Bone," In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. III, N. Y. Academic Press Inc., 1973
3. Pritchard, J. J., "General Histology of Bone," In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. I., N.Y. Academic Press Inc., 1972
4. Boyde, A., "Scanning Electron Microscope Studies of Bone" In G. H. Bourne (ed.): The Biochemistry and Physiology of Bone, Vol. I., N.Y. Academic Press Inc., 1972
5. Cameron, D. A., "Ultrastructure of Bone", In G. H. Bourne (ed.): The Biochemistry and Physiology of bone, Vol. I, N. Y. Academic Press Inc., 1972
6. Jones, S. J., A. Boyde and J. B. Pauley, "Osteoblast and Collagen Orientation", Cell Tiss. Res., Vol. 159, pp. 73-80, 1975
7. Graham, M.D. and H. P. House, "Human Stapes Crura," Arch. Otolaryngol., Vol. 101, pp. 548-551, 1975
8. Hayat, M.A. Introduction to Biological Scanning Electron Microscopy, Baltimore, University Park Press, 1978.
9. Peck, W. A., J. K. Burks, J. Wilkins, S. B. Rodan, and G. A. Rodan, "Evidence of Preferential Effects of Parathyroid Hormone, Calcitonin and Adenosine on Bone and Periosteum," Vol. 100, pp. 1357-1364, 1977.
10. Whedon. G. D., Lutwak, L., Reid, J., Rambaut, P. Whittle, M., Smith M. and Leach, C., Mineral and Nitrogen Metabolic Studies on Skylab Orbital Space Flight, Trans. Assoc. Am. Phys. 87:95, 1976
11. Bassett, L. S., Tzitikalakis, G., Pawluk, R. J., and Basset, C. A., Prevention of Disuse Osteoporosis in Rats by Means of Pulsing Electromagnetic Fields, In C.T. Brighton (ed.): Electrical Properties of Bone and Cartilage, Grune and Stratton, New York, 1979.
12. Birge, S. J. and Whedon, G. B., Bone, In U. MacNally (ed.): Hypodynamics and Hypogravics, Academic Press, New York, 1968.
13. Rajiv, K. and Riggs, B. L., Pathologic Bone Physiology, In M. R., Urist (ed.): Fundamental and Clinical Bone Physiology, J. B. Lippincott Co., Philadelphia, 1980.



14. Kazarian, L. E. and von Gierke, H. E., Bone, Loss as a Result of Immobilization and Chelation: Preliminary results in Macaca mulatta, Clin. Orthop., 65:65-67, 1969
15. Kazarian, L. E., Cann, C. Parfitt, M. Simmons D. and Morey-Holton, E., Al4-day ground-based hypokinesia study in nonhuman primates-a compilation of results. Nasa-TM-81268, 1981.
16. Wronski, T. S., Morey E. R., Inhibition of Cortical and Trabecular Bone Formation in Long Bones of Immobilized Monkey.
17. Morey, E. R. and Baylink, D. J., Inhibition of Bone Formation During Space Flight, Science, 201, 1138-1141, 1978.
18. Humphreys, Walter J., B. O. Spurlock and J. S. Johnson. "Critical Point Drying of Ethanol-Imfiltrated Cayofracted Biological Specimens for Scanning Electron Microscopy", Scanning Electron Microscopy/ 1974 (Part I) Proceedings of the SEventh Annual Scanning Electron Microscope Symposium, IIT Research Institute, Chicago, Illinois, 275-282.
19. Binderman, Itzhak, D. Duksin, A. Harell, E. Katzir and Leo Sachs, Formation of Bone Tissue in Culture from Isolated Cells, The Journal of Cell Biology, 61:427-439, 1974.
20. Yagiela, John A and D. M. Woodbury, Enzymatic Isolation of Osteoblast from Fetal Rat Calvaria, Anat. Rec., 188:287-308, 1977.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ON A THIN-LAYER NAVIER-STOKES CODE AND TRANSONIC

PROJECTILE AERODYNAMICS

Prepared by: Chen-Chi Hsu

Academic Rank: Professor

Dept and University: Department of Engineering Sciences  
University of Florida

Research Location: Air Force Armament Laboratory  
Aeromechanics Division  
Aerodynamics Branch

USAF Research: Dr. Lawrence Lijewski

Date: 1 August 1984

Contract No: F49620-82-C-0035

ON A THIN-LAYER NAVIER-STOKES CODE AND  
TRANSONIC PROJECTILE AERODYNAMICS

by

Chen-Chi Hsu

ABSTRACT

An Axisymmetric thin-layer Navier-Stokes code and a grid generation<sup>code</sup> have been installed and studied for their application, implication and effectiveness to the computation of transonic projectile aerodynamics. Preliminary numerical results obtained for four different flow cases,  $M = 0.91, 0.94, 0.96$ , and  $0.98$ , seem to indicate that the thin-layer Navier-Stokes code can provide satisfactory results if a good adaptive grid network is used in the computation. It clearly indicates the importance of developing an automatic adaptive grid generation code for use in the Navier-Stokes code. Moreover, additional numerical experiments are required to assess the accuracy of the algebraic turbulent model programmed in the code for the flow characteristics downstream of a shock.

## I. INTRODUCTION

An accurate prediction of the aerodynamic drag force is essential to a better design of aerodynamic devices and flight vehicles. In general, the total drag force can be divided into three distinct parts, that is, the surface pressure drag, the viscous drag and the base drag. For a transonic projectile aerodynamic problem, the relative magnitude of the drag forces is roughly about 20% for the pressure drag, 30% for the viscous drag, and 50% for the base drag. At present state of computer technology, an accurate numerical prediction of these drag forces for a complex flow problem is rather difficult and involved; in fact, the number of grid points which can be used in the simulation is often limited by the capacity of the existing computer systems.

Recently a thin-layer Navier-Stokes code had been developed for three-dimensional compressible fluid flow problems [1]. The code can provide unsteady or steady inviscid and viscous flow solutions; for the viscous case, one can further specify either laminar or turbulent flow. The turbulent closure model programmed in the code is a two-layer algebraic eddy viscosity model which is a modification of the Cebeci's model for avoiding the finding of the boundary-layer edge [2]. The Navier-Stokes code also had been simplified for axisymmetric flow problems to improve the computational effectiveness [3]. The Navier-Stokes codes had been shown to provide acceptably accurate solution for a number of flow problems.

The application of the thin-layer Navier-Stokes codes to transonic projectile aerodynamic problems has been investigated by the U.S. Army Ballistic Research Laboratory. For a SOCBT projectile at zero angle of attack, as reported in [4,5], the computed pressure coefficient  $C_p$  over the

secant-ogive portion and the boattail portion agrees rather well with the experimental data; however, the agreement on the cylinder portion of the projectile is generally not very satisfactory. For the projectile at  $2^\circ$  angle of attack and  $M = 0.91$ , the  $C_p$ -distributions computed from a CRAY 1S computer and a CDC system do agree qualitatively with the experimental data but quantitatively the computed  $C_p$  over the cylinder and boattail portions of the projectile is not satisfactory [6]. The unsatisfactory results computed can be attributed to the use of an improper grid network or to the thin-layer approximation of the Navier-Stokes equations as well as to the solution algorithm and its application.

## II. OBJECTIVE

It is believed, however, that the thin-layer Navier-Stokes code with necessary modification will have great potential applications to solving complex transonic projectile at high angle of attack, finned projectile, and other projectile related problems which are of interest to Air Force if a properly adaptive grid network is provided. Therefore, the main objective of this research is to assess the accuracy of the thin-layer approximation and the implication of the solution algorithm in the thin-layer Navier-Stokes code for transonic projectile aerodynamic computation. A secant-ogive-cylinder-boattail (SOCBT) projectile shape, a grid generation code, and an axisymmetric version of the thin-layer Navier-Stokes code obtained from the U.S. Army Ballistic Research Laboratory will be employed in this investigation.

### III. GRID NETWORK

A grid generation code named GRIDGEN had been developed at BRL for the computation of transonic projectile aerodynamics [7]. It can provide a good adaptive grid system for the transonic flow problem if the boundary grid points are properly chosen. The code can generate a two-dimensional elliptic grid network which is then modified with an exponential clustering along the grid lines normal to the streamwise direction to provide sufficient grid resolutions for the viscous region. For projectile aerodynamics, a three-dimensional grid network is formed by generating a sequence of planar grid networks about the axis of the projectile.

The GRIDGEN had then been implemented to have the capability of generating a two-dimensional hyperbolic grid network [8]. However, the option of hyperbolic grid generator had not been used in aerodynamic computation since the grid network generated had some undesirable characteristics in the region upstream of the projectile nose. In the summer of 1982, the hyperbolic grid generator was modified and improved to provide a competitive grid network for projectile aerodynamic computation; preliminary numerical experiments conducted showed that both elliptic grid network and hyperbolic grid network generated from modified GRIDGEN are about the same in effectiveness [9].

It is observed that the hyperbolic grid network generated has better overall characteristics than the elliptic grid network, in particular, the nearly orthogonal grids in the viscous region which is highly desirable to minimize the error. Therefore, the hyperbolic grid network is employed in the present study. A 78x28 hyperbolic grid network generated and used in this study is shown in Figure 1.

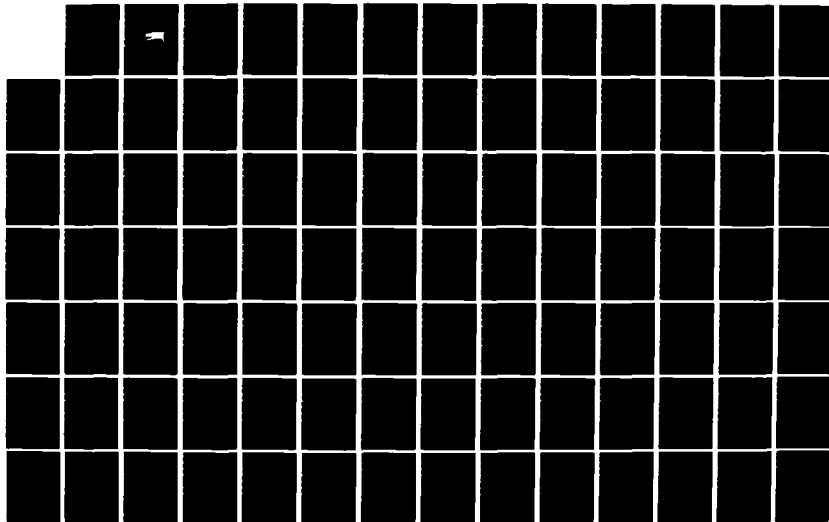
AD-A154 337 UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 5/13

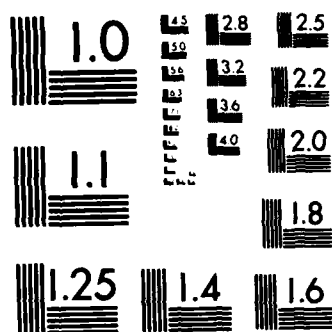
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

UNCLASSIFIED W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL

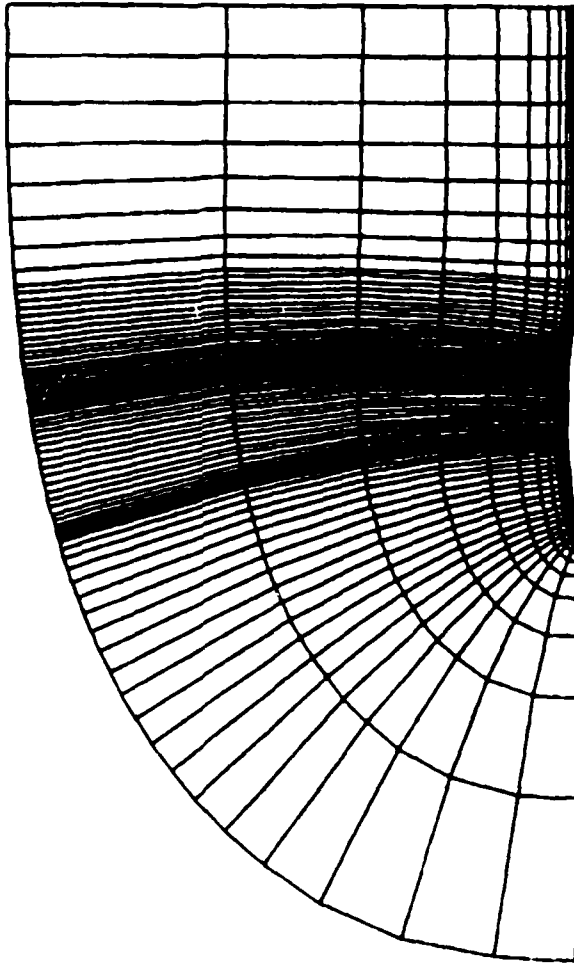




MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A



Figure 1. HGRID 78/28 ORIGINAL



#### IV. THIN-LAYER NAVIER-STOKES CODE

The axisymmetric version of three-dimensional thin-layer Navier-Stokes code obtained from the U.S. Army Ballistic Research Laboratory has been installed for this investigation on steady transonic projectile aerodynamics at zero angle of attack. The theoretical background of the code can be found in the reference [3]. It is mentioned in passing that the transformed governing equations of motion are simulated by Beam and Warming's factorized finite-difference scheme [10] in which a second order implicit and a fourth order explicit artificial dissipation terms have been added for controlling the numerical stability problem.

The documentation of the thin-layer Navier-Stokes code does not exist. The code apparently is a research code; consequently, in similarity to the grid generation code GRIDGEN, the dimension of some variables is not consistent with that of others. However, the use of the Navier-Stokes code is rather straightforward. First, the grid network generated from GRIDGEN must be provided in File or Tape 3. Then, the following six sets of formatted input data must be specified.

1. NMAX,JMAX,KMAX,LMAX,ND; format (5I5).

NMAX = maximum number of time steps to be performed.

JMAX = number of grid points in the streamwise ( $\xi$ ) direction.

KMAX = 1 for the axisymmetric case.

LMAX = number of grid points in the normal ( $\eta$ ) direction.

ND = 1 for the axisymmetric case.

2. DT,DX,DY,DZ,FSMACH,SMU,ALP; format (7F10.1).

DT = time step size to be used for NMAX integration.

DX = DY = DZ = 1.0 since they are irrelevant in the code.

FSMACH = freestream Mach number.

SMU = explicit smoothing (dissipation) parameter, .

ALP = 0. for axisymmetric flows.

3. IREAD,IWRIT,NGRI,INVISC,NP; format (5I5).

IREAD = { 0 for generating the initial conditions,  
1 for reading the initial conditions from File 2.

IWRIT = { 0 for not creating a restart file,  
1 for generating a restart file.

NGRI = { 1 for 2-D grid network in File 3,  
2 for 3-D grid network in File 3.

INVISC = { 0 for inviscid flow,  
1 for viscous flow.

NP = NP for printing results at every NP time steps.

4. CNBR,OMEGA,PDOV; format (3F10.0).

CNBR = Courant number. CNBR = 0. implies that Courant number will be

computed with specified DT. If CNBR is assigned then DT is determined from CNBR.

OMEGA = angular speed of the projectile in rpm.

PDOV = dimensionless velocity in O-direction.

5. RE,PR,RMUE,TZ; format (7F10.0).

RE = Reynolds number.

PR = Prandtl number.

RMUE = 1.0 for turbulent flow.

TZ = freestream stagnation temperature.

6. LAMIN,RM; format (I5, F10.0).

$$\text{LAMIN} = \begin{cases} 0 & \text{for laminar flow,} \\ 1 & \text{for turbulent flow.} \end{cases}$$

RM = implicit smoothing parameter, .

It has been recommended for steady solution to use in order DT = 0.005, 0.01 and 0.05 each for 50 steps and then use DT = 0.1 for the rest of computation.

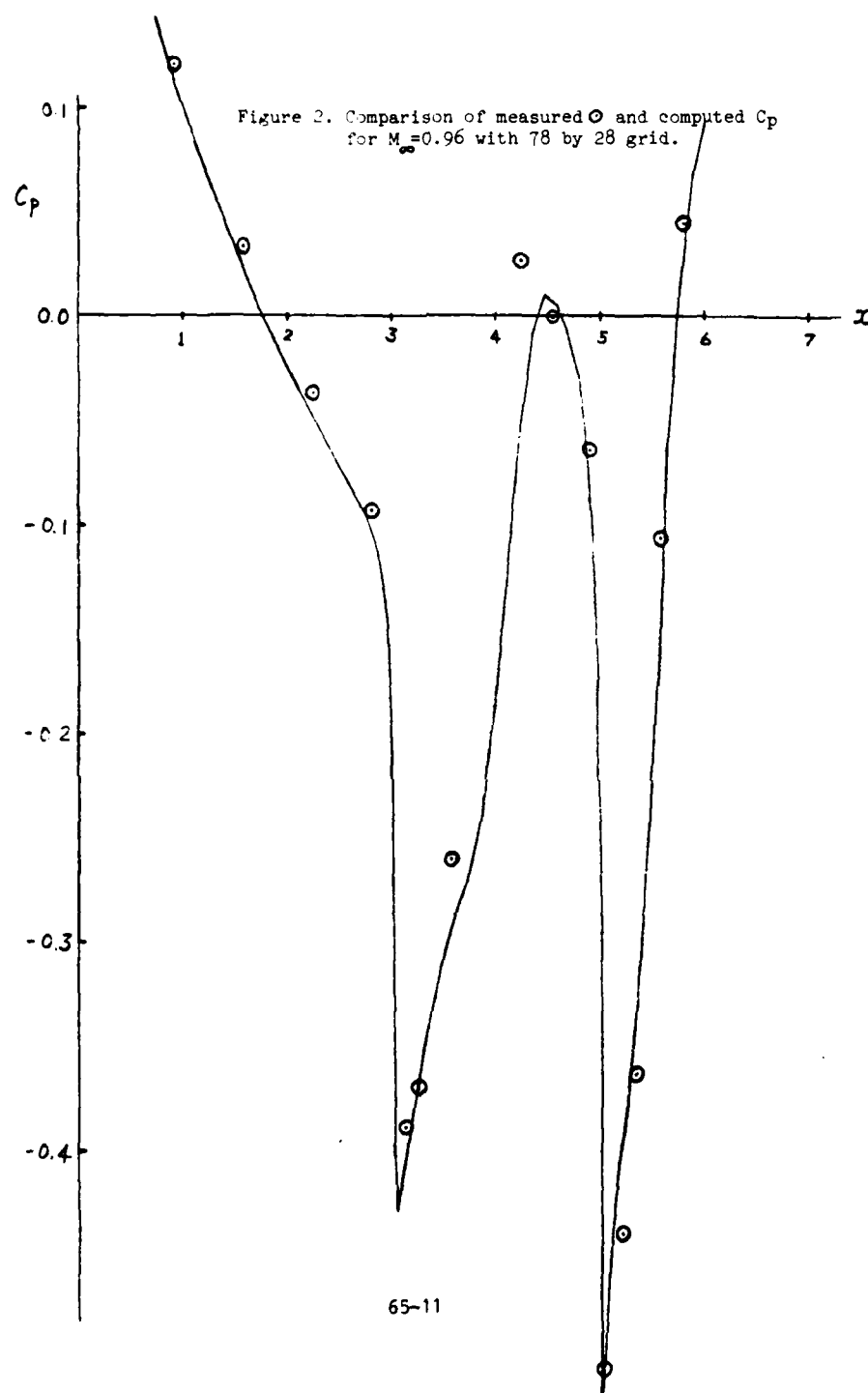
## V. RESULTS AND DISCUSSION

Some experimental data on the surface pressure coefficient had been reported for a 3-caliber secant-ogive nose, 2-caliber cylinder and 1-caliber 7° boattail shape projectile at transonic speeds [5]; hence, the projectile is chosen for present study. For the numerical experiment, however, the

projectile model used in the computation is the SOCBT projectile with the base region modified by a natural extension of the boattail for 1.75 calibers and then turned into a horizontal sting for 8.25 calibers.

The first grid network generated for use in the thin-layer Navier-Stokes code is a 60x28 hyperbolic grid system. The distribution of 60 grid points along the projectile model is 1-23-36-52-60 which implies that the number of discretization made for the secant-ogive portion, the cylinder portion, the modified boattail portion and the sting is 22,13,16 and 8, respectively. Furthermore, for each segment of the projectile, the grid points are adaptively distributed with a clustering function. A nearly converged solution has been obtained for the case of Mach number  $M_\infty=0.96$ . The computed pressure coefficient over the secant-ogive portion and part of the boattail agrees well with the experiment data; however, the agreement over the cylinder portion is not satisfactory at all. It seems to imply that a finer grid resolution is needed, in particular, for the cylinder portion of the projectile.

The second grid network generated for use is, therefore, a 78x28 hyperbolic grid system shown in Figure 1. The distribution of the grid points along the projectile is 1-23-46-70-78 which indicates that the number of grids on cylinder and boattail has been increased by 10 and 8, respectively. A practically converged solution for the case of  $M_\infty=0.96$  had been obtained from the Navier-Stokes code. The computed pressure coefficient and the reported experimental data are shown in Figure 2 for comparison. Indeed the agreement is excellent for the secant-ogive and boattail portions and is acceptable for the cylinder portion. It is noted that the process of integration for a steady solution had failed at one time apparently due to a large correction in



the convergent process which resulted in taking the squared root of a negative number in the subroutine MUTUR of the thin-layer Navier-Stokes code. The difficulty can be overcome by either decreasing the size of integration time step  $\Delta t$  or increasing the value of artificial dissipation parameters  $\epsilon_x$  and  $\epsilon_y$ . The result of numerical experiments has shown that the doubling of  $\epsilon_x$  and  $\epsilon_y$  has little effect on a converged pressure distribution. The sufficiency of the adaptive boundary gridding for each segment of the projectile also has been verified by the application of two additional  $78 \times 28$  grid networks; the distribution of pressure coefficient computed with the two grid networks is almost right on the top of the distribution shown in Figure 2.

To further positively assert the effectiveness of the thin-layer Navier-Stokes code for a solution of the transonic projectile aerodynamic problem, a more refined  $90 \times 28$  hyperbolic grid network has been generated for use in the Navier-Stokes code. For this grid network the grid point distribution along the projectile is 1-25-58-82-90. It is clear that ten more grid points have been added to the cylinder portion in hoping for an improvement on the solution accuracy in that region. For the case of  $M_\infty = 0.96$  considered, a practically converged solution has been obtained from the Navier-Stokes code at 1000 integration time steps. As one would have expected the computed pressure coefficient, in comparison with the experimental data, is more accurate than that obtained from  $78 \times 28$  grid network and presented in Figure 2. In fact the characteristics of the computed pressure coefficient distribution over the cylinder portion is more conforming with the behaviour dictated by measured data, which was not observed in the computed result reported in [4].

It is of interest and importance from both physical and numerical

viewpoints to observe the development of a steady solution computed from the thin-layer Navier-Stokes code. For the case of  $M_\infty=0.96$  solved with the 90x28 grid network, the initial condition generated in the code to start the integration process is the freestream condition; hence, the results computed are solutions to an impulsive transient transonic projectile aerodynamic problem. The computed surface pressure coefficient  $C_p$  at a number of different time steps  $N$  has been plotted to provide some information on the development of transonic flow field and on the convergence of a steady solution. Figure 3 presents the computed  $C_p$ -distributions over the secant-ogive portion of the projectile; it shows that the solution over the secant-ogive nose seems to have converged to an acceptable accuracy at  $N=400$  which corresponds to the dimensionless time  $\tau=24.5$ . However, the  $C_p$ -distributions over the cylinder and boattail segments shown in Figure 4 clearly indicate that a converged solution of the flow problem takes place at a much later time  $\tau=84.5$ , i.e., at  $N=1000$ . The transient  $C_p$ -distributions shown in Figure 4 seem to indicate that the shock at the ogive-cylinder juncture and at the cylinder-boattail juncture is nearly developed at  $N=300$  ( $\tau=14.5$ ). It is observed that the interaction of a developed shock and the flow field downstream does take a little while longer to come to steady; moreover, the converging process of  $C_p$ -distributions shown in Figures 3 and 4 implies that the interaction between a nearly developed shock and the flow upstream is rather minimal. Finally, one observes that the agreement between the converged  $C_p$ -distribution and the experimental is excellent except for a small region around the middle of the cylinder.

The 90x28 hyperbolic grid network has been employed again for solving three additional transonic flow problems at  $M_\infty=0.91$ , 0.94 and 0.98, respectively. It was reported in [4] that the number of integration steps



Figure 3.  $C_p$ -distribution over secant-ogive portion  
for  $M_\infty=0.96$  at various time step  $N$ .

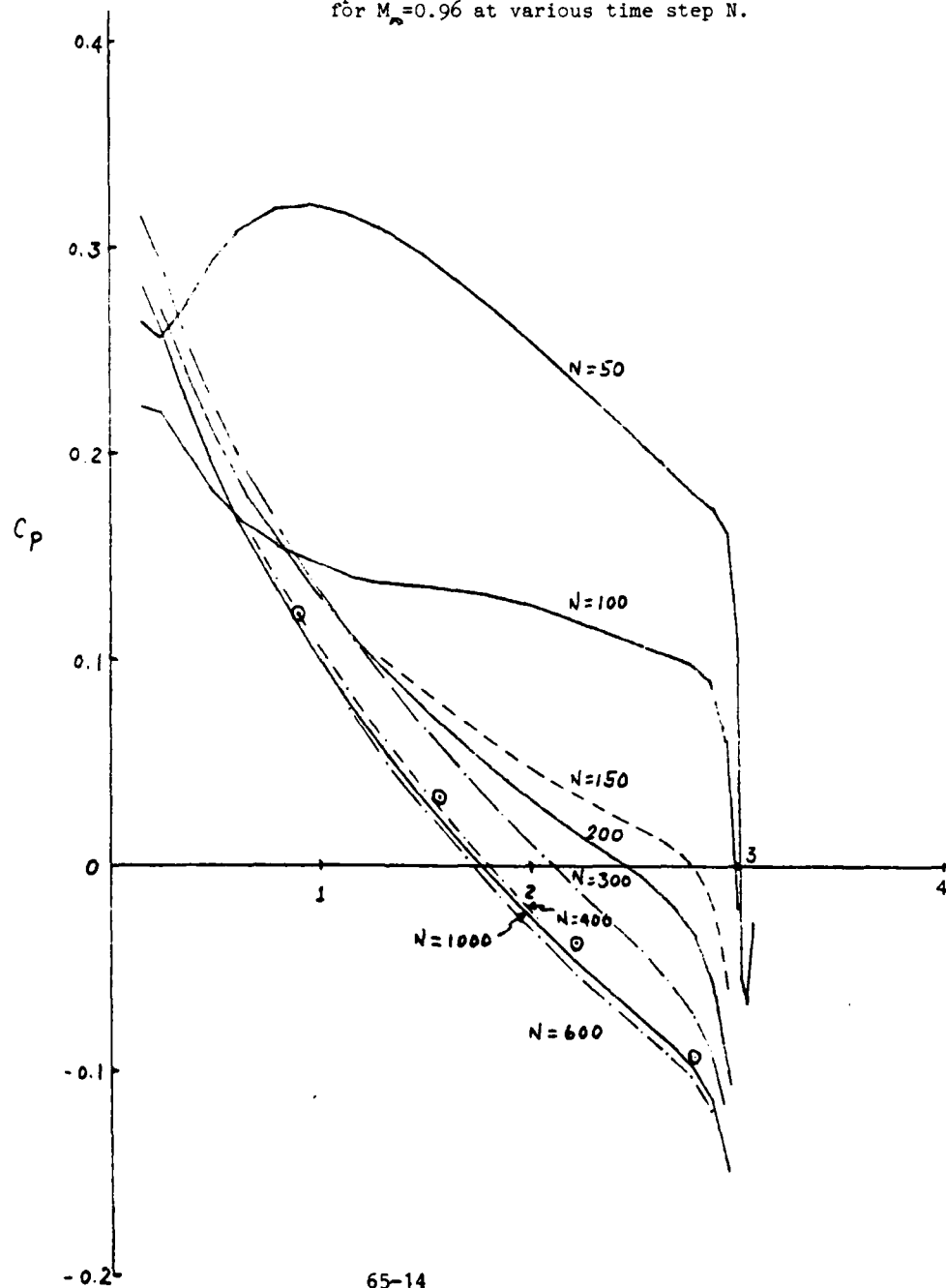
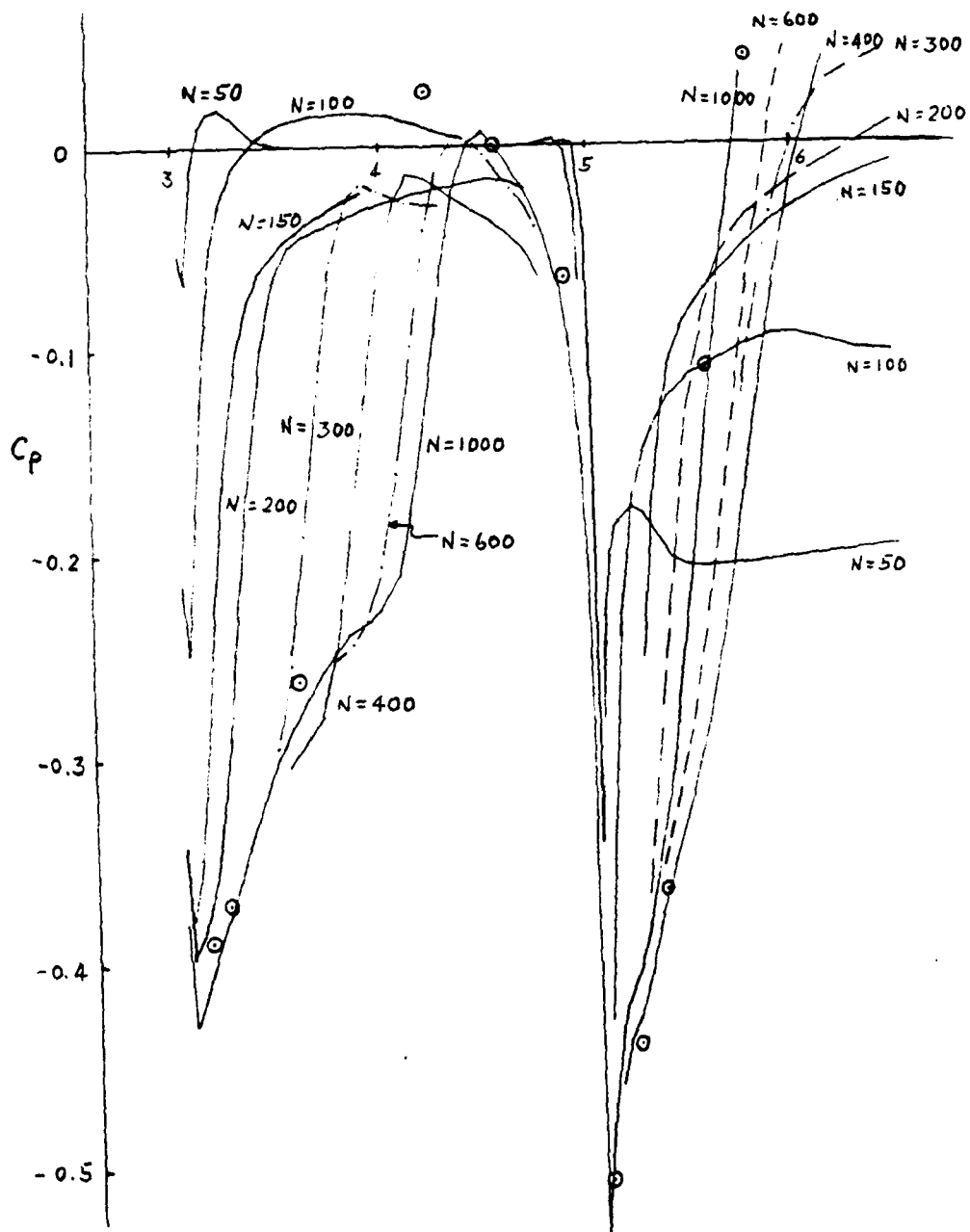


Figure 4.  $C_p$ -distribution over cylinder and boattail portions  
for  $M_\infty=0.96$  at various time step  $N$ .



required to achieve a steady solution can be reduced significantly if a converged solution of a similar problem is chosen as the initial condition to start the integration process. Hence, for each of the three additional flow problems considered, a selected intermediate solution of the case  $M_\infty=0.96$  has been used in the Navier-Stokes code to start the integration process; however, contradicting with the finding reported in [4], a practically converged solution can be obtained only after about 1000 integration steps. The results obtained for the three cases are very surprising in the sense that the  $C_p$ -distributions are nearly the same as that of  $M_\infty=0.96$ . It seems to imply that the solution computed for the three additional flow problems is trying to converge to the steady solution of  $M_\infty=0.96$  case. The measured data for the four different flow problems reported in [5] show that the  $C_p$ -distributions over the secant-ogive nose upstream of a shock are very similar to each others; however, the  $C_p$ -distribution downstream of a shock depends rather strongly on the Mach number. In fact the characteristics of the  $C_p$ -distribution measured for  $M_\infty=0.91$  is quite different from that measured for  $M_\infty=0.98$  over the cylinder portion of the projectile.

All of the facts indicate that the computed results for the three additional flow problems are not correct and that there are errors either in the input data or in the solution algorithm. After some thought and searching as well as very careful examination of the input data and printout results the source of errors has been located. It is found that a solution of another problem in general cannot be used as the initial condition to start the new problem in the thin-layer Navier-Stokes code without modifying the code. In fact, with the correct input data for a new problem, the subroutine INITIA of the Navier-Stokes code computes the initial condition at every point of the grid network based on the free stream condition. However, when a restart file

is read, the originally computed value for all the dependent variables at every grid point is overridden by the data in the restart file. Therefore, the correct outer boundary conditions of the new problem are replaced by those of the other problem for the subsequent integration process. This fact explains why the computed results for the three additional flow problems were trying to converge to the solution of the flow case  $M_\infty=0.96$ . It is mentioned in passing that a simple modification of the subroutine INITIA can overcome the difficulty so that the Navier-Stokes code can accept a solution of another problem as the initial condition and yet provides correct steady solution for the new problem.

For the steady  $C_p$ -distribution computed and presented in Figures 3 and 4 for the flow case of  $M_\infty=0.96$ , the causes for the appreciable discrepancy between the measured data and computed  $C_p$  over a small portion of the cylinder are yet to be investigated. The discrepancy can be resulted from a number of sources such as insufficient grid resolution, improper boundary conditions at the chosen outer boundary or the turbulent model. A close examination of solution profiles at two streamwise stations indicates that the outer boundary chosen for free stream conditions may not be far enough from the body. For investigating the importance of the boundary condition a new 90x40 hyperbolic grid network has been generated. The outer boundary of the grid is about four projectile lengths from the body, which is an increase of about 33% from the old grid, while the grid point distribution along the projectile model remains the same as that of the 90x28 grid network. The first problem solved with the new grid network is the flow case of  $M_\infty=0.91$ . A preliminary analysis of results obtained from the Navier-Stokes code indicates that the predicted  $C_p$ -distribution over the cylinder portion does exhibit correct characteristics as dictated by the measured data; however, a nearly converged solution again

misses the measured peak  $C_p$  just downstream of the shock with the discrepancy in magnitude similar to that shown in Figure 4. The other three flow cases  $M_\infty=0.94$ , 0.96 and 0.98 are being solved with the 90x40 grid network. A detailed analysis of the computed results will be carried out after the completion of the SFRP.

## VI. CONCLUDING REMARKS

A rather brief study carried out on the derivation of the thin-layer Navier-Stokes equations and the solution algorithm and structure of the Navier-Stokes

code has made it possible to promptly identify and resolve a few difficulties encountered in the course of numerical investigation. It is noted that both the grid generation code and Navier-Stokes code appear to be in the form of a research code; consequently, there are unnecessary executable statements remaining in the codes and the dimension specified may not be consistent from one variable to another. It is mentioned in passing that the subroutine MAP in the Navier-Stokes code is not working properly at the present set up.

The numerical experiments conducted and the results obtained and analyzed have indicated that the thin-layer Navier-Stokes code can give rather satisfactory results for transonic projectile flow problems if a properly adaptive grid network is used in the computation. For further positive assertion, however, the results of the last set of numerical experiments must be carefully analyzed; moreover, a better grid network adaptive to the peak  $C_p$  region on the cylinder portion must be generated and used for the four flow cases. Then one can definitely justify the accuracy of the algebraic

turbulent model programmed in the code for flow characteristics just downstream of the shock.

It is recommended to develop a truly adaptive grid generation code for use in the Navier-Stokes code; it may be the key to solve complex projectile flow problems successfully. It is also recommended to assess the application and effectiveness of a three-dimensional thin-layer Navier-Stokes code for transonic projectile aerodynamics. The completion of the recommended work should provide sound foundation for tackling complex projectile flow problems such as stores hanging under the wing of a plane, which is of extreme interest to the Air Force.

#### ACKNOWLEDGEMENTS

This research was sponsored by the U.S. Air Force Systems Command, Air Force Office of Scientific Research. The cooperation and hospitality extended by members of the Aerodynamics Branch, Armament Laboratory, Eglin AFB are much appreciated. It is acknowledged that all the computations were performed by CRAY 1S computer system at NASA Ames Research Center through a terminal at Eglin AFB. This report was typed by Lisa Brown, AFATL/DLCA.

#### REFERENCES

1. Pulliam, T.H. and Steger, J.L., "Implicit Finite-Difference Simulations of Three-Dimensional Compressible Flow," AIAA Journal, Vol. 18, Feb. 1980.
2. Baldwin, B.S. and Lomax, H., "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," Paper 78-257 presented at AIAA 16th Aerospace Sciences Meeting, Huntsville, Alabama, Jan. 1978.
3. Nietubicz, C.J., Pulliam, T.H. and Steger, J.L., "Numerical Solution of the Azimuthal-Invariant Thin-Layer Navier-Stokes Equations," Paper 79-0010 presented at 17th Aerospace Sciences Meeting, New Orleans, LA, Jan. 1979.
4. Nietubicz, C.J., "Navier-Stokes Computations for Conventional and Hollow Projectile Shapes at Transonic Velocities," AIAA-81-1262, 14th Fluid and Plasma Dynamics Conference, Palo Alto, CA, June 1981.
5. Kayse, L.D. and Whiton, F., "Surface Pressure Measurements on a Boattailed Projectile Shape at Transonic Speeds," ARBRL-MR-03161, ADA 113520, U.S. Army Ballistic Research Laboratory, March 1982.
6. Sturek, W.B., "Opportunities for Application of Adaptive Grids in Computational Aerodynamics," in Adaptive Computational Methods for Partial Differential Equations (Proceedings), edited by I. Babuska et al., SIAM, 1983.
7. Steger, J.L., Nietubicz, C.J. and Heavey, K.R., "A General Curvilinear Grid Generation Program for Projectile Configurations," ARBRL-MR-03142, U.S. Army Ballistic Research Laboratory, Oct. 1981.
8. Nietubicz, C.J., Heavey, K.R. and Steger, J.L., "Grid Generation Techniques for Projectile Configurations," U.S. Army Ballistic Research Laboratory.
9. Hsu, C.C., "Grid Network Generation with Adaptive Boundary Points for Projectiles at Transonic Speeds," ARBRL-TR-02485, U.S. Army Ballistic Research Laboratory, April 1983.
10. Beam, R.M. and Warming, R.F., "An Implicit Finite-Difference Algorithm for Hyperbolic Systems in Conservation-Law Form," J. Comp. Physics, Vol. 22, 1976.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

EFFECT OF DISPLAY DYNAMICS IN MANUAL CONTROL TASKS

Prepared by:	Dr. Mario Innocenti
Academic Rank:	Assistant Professor
Department and University:	Department of Aerospace Engineering Auburn University
Research Location:	Air Force Wright Aeronautical Laboratory, Flight Dynamics Laboratory, Flight Control Division, Control Dynamics Branch
USAF Research	Frank L. George
Date:	August 17, 1984
Contract No:	F49620-82-C-0035



# EFFECT OF DISPLAY DYNAMICS IN MANUAL CONTROL TASKS

by

Mario Innocenti

## ABSTRACT

New advanced displays have shown to alter the flying qualities of aircraft. Present handling quality-specifications do not include the effect of display when the manual control task is performed. The present research performs a preliminary of the influence of display on a single axis pursuit tracking task. The display is modeled by a set of parameters and the relation between these parameters, tracking performance and pilot ratings is established using a fixed base simulation. Suggestions for further research on this area are offered.

## I INTRODUCTION

The understanding of the interaction between the components involved in a manual control task has become very important in recent years among the aircraft flying qualities community. This is especially true if we consider the advances in particular areas of aircraft technology, such as avionics, computer generated displays, digital controls, etc. These improvements, coupled with the development of new, highly maneuverable unconventional flight vehicles have led to the necessity of a new and more comprehensive approach to the flying qualities evaluation problem. It is shown, in fact, that the display dynamics play a primary role in influencing the pilot's judgement of the total system's capabilities since the pilot is flying certain dynamics, which consist of an integration of flight vehicle, controls, sensors, command and displays.

With the advent of new high performance systems capable of performing unconventional and untested tasks, it is necessary for the designer to furnish the pilot with a displayed information that is accurate for the task and helpful for the success of the mission. The questions that arise in this context are, therefore, not only what to display, but how to display it, how much information is needed and what kind of coupling exists among all the components of the system when the pilot closes the control loop. The long term objectives of this type of research are, eventually, to be able to establish some analytical criteria for flying qualities evaluation for the case in which the manual control problem is considered as an integrated process involving not only traditional components such as aircraft and pilot's dynamics, but display dynamics as well.

At this point, a brief review of the work done in recent years is necessary. Display analysis and design has dramatically evolved in recent years because of availability of sophistication reached by computer graphics techniques and digital computers. Still, display design consists largely of empirical procedures based on the application of extensive simulation. This interactive procedure is usually time consuming and very expensive. Clearly, simulation and experimental evaluation are necessary and unavoidable, but a need for some general analytical guidelines that would establish desirable characteristics and help the designer in a preliminary screening evaluation, has been suggested and recommended. The effect of compensatory versus pursuit display has been analyzed since the mid 50's early 60's by Elkind, McRuer and others<sup>1, 2, 3</sup>. Since then, a lot of work has been done in the general area of display design. From the human factors standpoint, recent studies by Roscoe<sup>4</sup> have analyzed improvement of display presentation using methods of quickening (rate feedback signal to give immediate presentation of imminent results) and prediction (estimation of future states based on present states). Jensen<sup>5</sup> has studied the same problem in a landing task and identified an evaluation based on performance. From the flying qualities standpoint, the bulk of the analytical work involves the use of models for the human operator.

McRuer and others<sup>6, 7</sup> used the frequency domain approach (i.e. crossover model) to gain insight into the effect of flight directors in the display design. The figure of merit was based on measured pilot workload and a landing approach was used as a test task. Time domain techniques have also been used (i.e. optimal control pilot model, or OCM) to analyze displays. Baron<sup>8</sup> and Curry<sup>9</sup> use performance and workload ratings from the analytical pilot model to rank the variables needed for display in a given task and to compare the influence of flight directors. Korn<sup>10</sup> applied the same modeling technique to evaluate four different displays in a Terrain Following task; Lebac<sup>11</sup> used flight test data for ranking displays with director in the presence of wind gusts. Finally, Hess<sup>12</sup> also applied OCM and provided a semi-analytical guideline for display design during to an instrument landing approach.

All the work mentioned above, deals mainly with one aspect of display design, which only marginally addresses the coupling problem between display dynamics and other dynamics in the manual control loop. Work more closely related to this area has been done by Hess<sup>13</sup> and Lowe<sup>14</sup>. Here, director and display dynamics were introduced and they were shown to influence the pilot's performance and ratings of the task. Finally, a preliminary work by Weener<sup>15</sup> is worth mentioning: in his experiment a relation between altitude tracking performance and the bandwidth of a filter interposed between true aircraft altitude and displayed altitude was shown.

From the bibliography review, some conclusions can be drawn. Over the years it has been recognized the need of a more analytical approach to the display design problem. Models of human operator in the control loop have been used in order to rank different display formats, this was done, mainly, in terms of pilot's workload and performance. Only relatively little has been done to investigate display dynamics-aircraft-pilot interaction and mostly from an exploratory and preliminary analysis standpoint.

The main interest of the Air Force, as well as other parties in the aerospace community is to investigate the possibility of defining some criteria that relate display dynamic properties, flying qualities and task performance requirements. In this respect, very little is available today especially in terms of pertinent mathematical models of the static and dynamic coupling between display, pilot, vehicle and environment. The present research effort is devoted to a preliminary investigation of such problems.

## **II RESEARCH OBJECTIVES**

The area of manual control of flight vehicles has always been of primary importance for the Air Force. Even more so in recent years with the advent of highly maneuverable, complex aircraft which present the pilot with new situations. In order to reach a complete understanding of the

problem and ultimately obtain guidelines and specifics for flying qualities, all the areas of manual control must be carefully analyzed.

Generally speaking, the manual control includes four major components: the aircraft or plant to be controlled, the information system or display which gives the pilot all the cues needed for performing his/her task, the commands and the human operator itself closing the loop. All these components are operating and interacting together with the objective of successfully accomplishing the mission. In this effort, we will focus our attention onto one of the components in particular, the display.

Display design has gone through a lot of changes with the introduction of advanced computer graphics capabilities and availability of faster and bigger on-board digital computers for signal processing. As a result of this development in technology, present and future aircraft rely and will rely on sophisticated computer generated displays such as head up and head down displays for monitoring as well as control during the various phases of the flight.

Two factors in the display design have become very important and need to be mentioned at this point. The first is the type of display. Due to the availability of generating different displays, the questions are now what do we display? How much information is enough? Most of the work described in the introduction addresses the solution of this problem. Experimental analysis and predictive tools using analytical pilot models are used to rate different displays in terms of pilot's workload and performance. The second aspect is more related to the area of flying qualities. It deals with the interaction of the display dynamics and the other elements in the manual control loop and it is critical in the determination and evaluation of the overall aircraft flying qualities.

Relatively little work has been done in this area, yet the effect of display dynamics plays a primary role in how the pilot rates the overall system. The present work attempts to perform a preliminary analysis of the effect of display dynamics on the manual control loop. The long term objectives of the present research could be envisioned as the development of analytical procedures, which will enable the designer to evaluate flying qualities of aircraft that include advanced displays. In particular, evaluation of displays, which include some form of state prediction, could be considered for tasks such as terrain following - terrain avoidance.

In order to begin the study, the present effort is organized as follows:

1. Definition of the display characteristics.
2. Definition of the dynamics of the manual control loop components and task.
3. Experimental evaluation and analysis.
4. Results and conclusion.

### III DISPLAY CHARACTERISTICS

In this section we will evaluate some of the parameters that can be used as representative of the display dynamics. The display is a device whose function is to give the pilot an accurate, continuous description of one of more independent variables of the system. Since the display is a dynamic system in itself, it is reasonable to assume that it will affect the global performance when the pilot closes the loop. Some work to validate the above consideration can be found in Reference 15, where it is shown that the display would not only change the pilot's performance, but also it would give origin to a dynamic coupling with the aircraft.

As for any dynamic system, the mathematical modeling of the display is probably the most difficult task, this difficulty is compounded by the fact that there is not a single or unique display. To this end, we will assume that the display is adequately represented by a set of parameters that appear to be common among different displays. Due to time limitations, we will only consider some of these parameters; they are:

1. Time Delay  $\tau_D$  that takes into account the processing time in the display. (Computation time, refresh rate, etc.).
2. Frequency  $\omega_D$ , which gives an indication of the display bandwidth, including the presence of prefilters.
3. Damping Factor  $\xi_D$ , which is related to the stability of the image and the time constant of the system.
4. Gain  $K_D$  relating the display input and output.

It must be kept in mind, however, that the above list is by no means exhaustive, further research is needed to obtain a more complete picture. In summary, the "generic" display considered herein can be represented schematically by the block diagram in Figure 1.

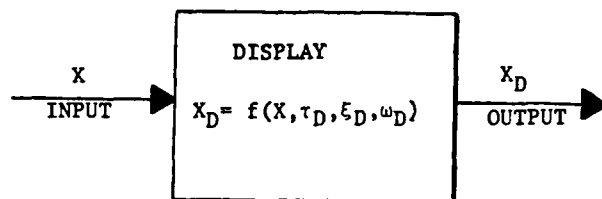


Figure 1. Display Schematic

With the above assumption in mind, our objectives are to establish a relationship between the display parameters and the performance of the manual control loop as a first step toward the identification of an analytical procedure, which would relate the dynamics of the system (including display) to flying qualities characteristics.

#### IV TASK AND SYSTEM STRUCTURE

As a preliminary analysis, a generic single axis tracking task is considered. Although this task is not fully representative of a multi-axis complex situation, it is simple enough that it can be analyzed in the time frame available and it offers insight into the physics of the problem.

A tracking task is, usually, of two types: compensatory and pursuit. In a compensatory tracking situation one single moving index is shown to the pilot, the direction and the distance from a fixed reference represents the tracking error to be nulled by the pilot. In a pursuit tracking environment, the moving indices are two instead of one, one represents the target motion and the other the actual aircraft position. The difference between the two cases is in that the compensatory task gives only one independent information, the tracking error, while a pursuit situation, allows the pilot to have two independent references among the three parameters (input, output, error) displayed. There are advantages and disadvantages to both cases, although it appears that pursuit is preferred<sup>2, 17</sup>. It must be noted<sup>3</sup>, however, that there are cases in which the pilot although presented with a pursuit type of situation, behaves in a compensatory fashion.

In this work, a pursuit display will be used since it appears to offer more flexibility and also it approximates more closely head down displays where the visual cues are absent and the pilot is flying an instrument task.

The manual control loop system structure is schematically shown in Figure 2.

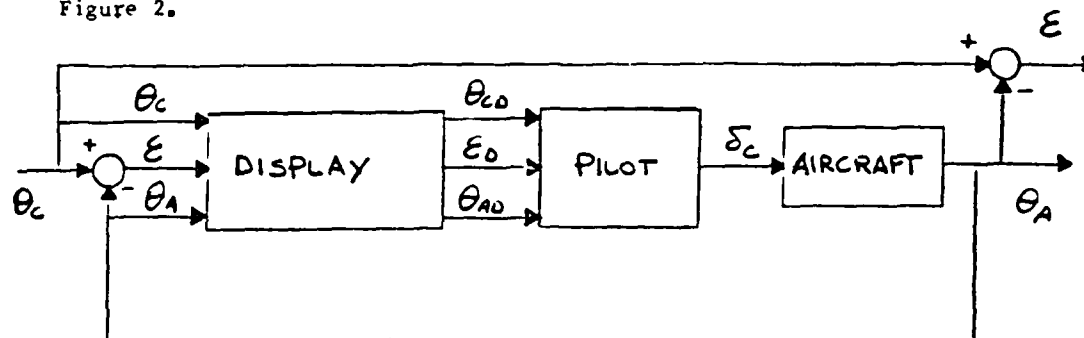


Figure 2. Manual Control Loop

The tracking signal  $\theta_c$  and the aircraft position  $\theta_A$  are displayed to the pilot. The pilot can extract the tracking error information  $\epsilon = \theta_c - \theta_A$ . The displayed variables become therefore  $\theta_{CD}$ ,  $\theta_{AD}$  and  $\epsilon_D$ . In order to model the display according to the parameters described in section III, we will assume that the dynamics of the display are given by the following transfer function

$$G_D(s) = G_F(s)G_T(s) \quad (1)$$

where

$$G_F(s) = \frac{K_D}{s^2 + 2\zeta_D \omega_D s + \omega_D^2} \quad (2)$$

and

$$G_T(s) = \frac{-(s - 2/\tau_D)}{(s + 2/\tau_D)} \approx e^{-\tau_D s} \quad (3)$$

Equation (2) is a second order filter with variable parameters. Equation (3) represents the delay present in the display and it will be modeled as a first order Pade approximation. In conclusion, the dynamics of the display are represented by a third order system with transfer function  $G_D(S) = \frac{K_D}{s^2 + 2\xi_D s + \omega_D^2} e^{-s\tau_D}$ , which consists of a second order filter and a pure time delay. In order to evaluate different display dynamics, the parameters  $\tau_D, \xi_D, \omega_D$  are chosen as in Table a.

Table a. Display Parameters

$\tau_D$ (SEC)	.001	.125	
$\xi_D$	.2	.707	
$\omega_D$ (RAD/SEC)	4.	8.	12.

A total of 12 different display combinations is therefore available, ranging from an underdamped system with almost no delay and small bandwidth (relative to the pilot's capability), to a critically damped system with relative large delay and bandwidth. The gain  $K_D$  represents the sensitivity of the display and it has a nominal value of unity.

The second component is the pilot. The pilot's behavior will not be discussed here in terms of a mathematical model, rather we will evaluate his performance and ratings of the overall system. The pilot's input to the aircraft is given by a stick command  $\delta_C$  (the stick dynamics are negligible here).

The third component is the aircraft. The aircraft dynamics are modeled by the transfer function

$$G_A(s) = \frac{\theta_A}{\delta_C} = \frac{K_A a}{s(s+b)} \quad (4)$$

The three parameters appearing in (4) are chosen from Reference 12 and they are repeated in Table b below.



Table b. Aircraft Parameters

$K_A$ (VOLT/VOLT)	a	b
.586	40	40
1.17	1.	0.
2.15	1.	4.

Three configurations are considered here. The first corresponding to a K/s type of plant, the second to a double integrator and the third to a second order, marginally stable system. In total, 36 cases will be simulated.

The tracking signal  $\theta_C$  is taken from Reference 16 and it consists of a sum of 5 sine waves reproducing a random signal.

$$\theta_C = \sum_{i=1}^5 \sin \omega_i t$$

The numerical values for amplitude and frequency are given in Table c.

Table c. Input Parameters

$A_i$ (VOLTS)	$\omega_i$ (RAD/sec)
.566	.188
.566	.502
.566	1.88
.113	7.35
.113	15.02

#### V SIMULATION

The organization and actual run of the experiment have been the major portion of the present effort. A detailed description of the simulation can be found in Reference 18. Here we will briefly review some of the properties.

The hardware used for the simulation consists of a hybrid MODCOMP computer available at the ASD Computer Center of WPAFB. The display is obtained using an oscilloscope driven by analog signals  $\theta_C$  and  $\theta_A$ . The pilot's command is through a displacement stick with for aft movement. The stick sensitivity is set at 2.5°/volt. The maximum voltage in the system is  $\pm 10$  volts.

One pilot was used for the simulation. After proper training, the pilot flew a total of 42 configurations. 36 of these were derived from the previous section and 6 were added more with the display progressively out of the loop. The subject was to perform a tracking task trying to keep the tracking error zero. The simulation time was 100 seconds per configuration, so that an integer number of cycles of the input sine waves was available. Data was recorded on tape for further analysis. The next section presents some of the results.

#### VI ANALYSIS OF THE RESULTS

Several sets of data were taken during the simulation. The time histories  $\theta_C$ ,  $\theta_A$ ,  $\epsilon$ ,  $\theta_{CD}$ ,  $\theta_{AD}$ ,  $\epsilon_D$  and  $\delta_C$  were obtained on a strip chart recorder as well as digitally on tape with a sample rate of 20 samples

per second. The sample data was used for spectral analysis, in order to obtain experimental describing functions of the system. Statistical analysis was also performed on the sampled data and average, rms values and standard deviation for the all the variables were computed. The rating scheme for the pilot's evaluation was based on the four scales shown in Table d. The scales are all from 1 to 10; the first is a tracking performance evaluation, the second rates the system's response, the third is a measure of pilot's workload and the fourth is a global evaluation of the configuration, comparable, although not equal, to a Cooper-Harper rating scale. Additional information was obtained in the form of pilot's comments after each configuration run.

A complete document of the results and their analysis can be found in Reference 18. Here, we will present some of the results for one of the three configurations and we will briefly analyze them.

The configuration examined in detail is the one corresponding to a single integrator (first row in Table b).

The results are shown in terms of RMS displayed tracking error and pilot ratings. The various cases shown correspond to different values of the display parameters.

Table d. Rating Scales

PERFORMANCE (1)		RESPONSE CHARACTERISTICS (2)	
0	- Extremely easy to track with excellent precision	0	- Excellent, pure (i.e., no accidental excitation) response characteristics
1		1	
2	- Very easy to track with good precision	2	
3		3	
4	- Easy to track with fair precision	4	- Good, relatively pure response characteristics
5		5	- Fair, somewhat impure response characteristics
6	- Possible to track, inadequate precision	6	- Quite sensitive, sluggish, or uncomfortable in response
7	- Possible to track, imprecisely	7	
8	- Difficult to track, very difficult to track	8	- Extremely sensitive, sluggish, or uncontrollable in response
9	- Nearly uncontrollable	9	- Nearly uncontrollable
10	<input type="checkbox"/> Uncontrollable	10	<input type="checkbox"/> Uncontrollable
	<input type="checkbox"/> Not Applicable		<input type="checkbox"/> Not Applicable

DEMANDS ON PILOT (3)		HANDLING QUALITIES (4)	
0		0	
1		1	- Excellent
2	- Completely undemanding, very relaxed & comfortable	2	- Highly Desirable
3		3	
4	- Largely undemanding, relaxed	4	- Good, pleasant, very minor deficiencies
5	- Mildly demanding of pilot attention, skill, or effort	5	- Fair, some objectionable deficiencies
6	- Demanding of pilot attention, skill, or effort	6	
7	- Very demanding of pilot attention, skill, or effort	7	- Bad
8	- Completely demanding of pilot attn, skill, effort	8	- Very bad
9	- Nearly uncontrollable	9	- Nearly uncontrollable
10	<input type="checkbox"/> Uncontrollable	10	<input type="checkbox"/> Uncontrollable
	<input type="checkbox"/> Not Applicable		

Table e. Simulation Results

CASE	$\tau_D$	$\xi_D$	$\omega_D$	$K_D$ (volt/div.)	$\epsilon_D$ (RMS)	RATING			
						1	2	3	4
1	.001	.2	4	2	.252	4	5	6	6
4	.001	.2	12	.5	.025	2	3	5	5
6	.125	.2	8	.5	.061	4	5	6	5
10	.125	.2	12	1	.029	2	3	3	4
22	.125	.707	8	1	.064	2	4	5	4
27	.001	.707	8	1	.047	3	4	4	4
33	.001	.707	12	2	.023	2	2	3	3
36	.125	.707	12	1	.04	2	4	4	4
14	.001	.707	4	1	.195	7	7	7	7
16	.125	.2	4	2	.484	8	8	8	8
19	.001	.2	4	2	.312	8	8	8	8
34	.125	.707	4	1	.284	8	7	7	7
40	.125	.707	8	1	.048	4	4	5	4
41	.125	.707	8	1	.233	8	8	8	9
42	.125	.707	8	1	.434	8	8	8	8

Due to an error in the computer program, case 19 is the same as 1 and the case with  $\omega_D = 8$  rad/sec is missing, however this mistake can be used to our advantage. Table e shows 3 groups corresponding to 3 different results. The first group received better ratings than the other two in all the categories. In particular, the pilot judges group 1 to have better tracking performance and gave better overall ratings. Group 2, on the contrary, received poor ratings. The RMS tracking error shows the same trend (lower for group 1). The main difference between the two groups is in the display bandwidth  $\omega_D$ , cases with larger bandwidth gave better performance in terms of tracking error as well as pilot ratings. It is interesting to note the discrepancy in pilot's ratings between cases 1 and 19; the two cases are the same and yet case 19 was given a poorer rating although the RMS error was roughly the same. One possible explanation could be that the pilot had a different tolerance in the middle of the simulation than at the beginning (due to improper training).

Another comment deals with the results in group 3. Group 3 consists of 3 configurations with the same dynamic characteristics as case 22 in group 1. The difference is in the display dynamics, more exactly case 40 has  $\theta_A = \theta_{AD}$ , case 41 has  $\theta_C = \theta_{CD}$  and case 42 does not contain display dynamics at all. The ratings and the RMS error show that when the tracking signal is not filtered through the display or when the display dynamics are neglected, a decrease in performance is obtained (cases 41 and 42). As for case 40, the tracking error has the same magnitude as that of case 22. The pilot was able to do a better job but he did not like the configuration as much. From his comments for cases 22 and 40, we discover that he did not like the higher sensitivity of case 40, which was due to a lack of filter in his command.

Let us return now to group 1, which had the best ratings and performance. Looking at the overall ratings, the subject showed a slight preference for the last five cases. It appears that the difference is due to the damping factor  $\xi$ , better ratings were given to critically damped cases. The effect of time delay is not clearly shown in the ratings, although some difference can be seen in terms of tracking performance. For example, comparing cases 22 and 27, we can see that RMS tracking error is lower when the time delay is smaller; however, the pilot does not clearly rate accordingly. Comparing cases 33 and 36, instead, the beneficial effect of a smaller time delay can be seen in the tracking performance and the rating as well. Interesting enough, case 22, which appears to have the best behaved parameters, also received the best ratings and gave the best tracking performance.

Some conclusions can be drawn at this point. Of the three parameters representing the display dynamics, the bandwidth  $\omega_D$  is the most influential. Pilot's ratings of the same configuration change drastically with  $\omega_D$ , from difficult to track - bad configuration ( $\omega_D = 4$ ) to very easy to track - excellent configuration ( $\omega_D = 12$ ). A very large  $\omega$ ; however, does not improve ratings neither performance ( $\omega_D =$  in case 40, 41, 42). The second most influential parameter appears to be the damping  $\xi_D$ . As far as time delay is concerned, some discrepancy exists between tracking performance (always better at low  $\tau_D$ ) and subject ratings. A possible suggestion is to increase  $\tau_D$  from 125 msec to larger values.

## VII RECOMMENDATIONS

The research effort described in the present report consisted of a preliminary analysis of the effect of display dynamics on handling qualities. A very simple laboratory task was performed and the results have shown that display parameters greatly affect the pilot's evaluation. In particular, display bandwidth was found to be the primary factor.

Several potential extensions can be envisioned at this point, that would lead to a better understanding of the problem.

- Apply the proposed procedure to different aircraft dynamics

- Reevaluate the effect of time delay since the results do not appear to be conclusive.
- Increase the number of parameters used to describe the display dynamics. An example is the inclusion of sampling rate (related to the smoothness of the image).
- Include in the display some form of dynamic prediction of the displayed variables.
- Apply the analysis to a more complex tasks such as terrain following-terrain avoidance.

Once satisfactory ranges for the display parameters are identified in terms of good flying qualities, then the display design could be treated as a conventional control problem in that the display characteristics could be changed so that the values of the forementioned parameters would be within the prescribed boundaries.

These recommendations can be envisioned as a follow-on research, which would provide benefits to the Air Force, especially in the areas of flying qualities of complex flight vehicles.

#### ACKNOWLEDGMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very fruitful summer at the Air Force Wright Aeronautical Laboratories, WPAFB, Ohio. He would also like to acknowledge the Flight Dynamics Laboratory and the Flight Control Division, in particular, the Control Dynamics Branch for their hospitality.

Finally he would like to thank Mr. Frank L. George for his collaboration and guidance, Mr. R.O. Anderson for the many helpful discussions and Mr. John R. Carter of ASD, Simulation Branch, for the organization and preparation of the experimental work. A particular thanks goes to U.S. Air Force Capt Paul Weaver for his help as a pilot in the simulation.

#### REFERENCES

1. McRuer D.T., Graham D., Krendel E., Riesener W. Jr., "Human Pilot Dynamics in Compensatory Systems", AFFDL-TR-65-15, 1965.
2. McRuer D.T., Krendel E., "Dynamic Response of Human Operators", WADC-TR-56-524, October 1957.
3. Wasicko R.J., McRuer D.T., Magdaleno R.E., "Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays", AFFDL-TR-66-137, December 1966.
4. Roscoe S.N., Carl L., Jensen R.S., "Flight Display Dynamics Revisited", Human Factors, 1981, 23(3), pp. 341-353.
5. Jensen R.S., "Prediction and Quickening in Perspective Flight Displays for Curved Landing Approaches", Human Factors, 1981, 23(3), pp. 355-363.
6. Clement W.F., McRuer D.T., Klein R.H., "Systematic Manual Control Display Design", 13th AGARD Guidance and Control Panel Symposium, October 19-21, 1971.
7. Weir D.H., Klein R.H., McRuer D.T., "Principles for the Design of Advanced Flight Director Systems based on the Theory of Manual Control Displays", NASA CR-1748, March 1971.
8. Baron S., Levison W.H., "An Optimal Methodology for Analyzing the Effects of Display Parameters on Performance and Workload in Manual Flight Control", IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-5, No. 4, July 1975.
9. Curry R.E., Kleinman D.L., Hoffman W.C., "A Design Procedure for Control/Display Dynamics", Human Factors, 1977, 19(5), pp. 421-436.
10. Korn J., Gully S.W., Kleinman D.L., "Validation of an Advanced Cockpit Display Design Methodology Via Workload/ Monitoring Tradeoff Analysis", 18th Annual Conference on Manual Control,



Dayton, Ohio, 1982.

11. Lebacqz J.V., Aiken E.W., "Experimental Investigation of Control-Display Requirements for VTOL Instrument Transition", AIAA, Journal of Guidance and Control, Vol. 1, No. 4, July-August 1978.
12. Hess R.A., "Analytical Display Design for Flight Tasks conducted under Instrument Meteorological Conditions", NASA-TMX-73146, August 1976.
13. Hess R.A., "Aircraft Control-Display Analysis and Design Using the Optimal Control Model of the Human Pilot", IEEE, Trans on Systems, Man and Cybernetics, Vol. SMC-11, No. 7, July 1981.
14. Lowe J.R., Ornelas J.R., "Application of Head-up Displays in Commercial Transport Aircraft", AIAA Journal of Guidance and Control, Vol. 6, No. 2, March-April 1983.
15. Weener E.F., Howe R.M., Pew R.W., "Effect of Visual Flight Display Dynamics on Altitude Tracking Performance in a Flight Simulator", 9th Annual Conference on Manual Control, 1973.
16. McDonnell J.D., "Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities", AFFDL-TR-68-76, January 1969.
17. Reid L.D., "The Measurement of Human Pilot Dynamics in a Pursuit-Plus-Disturbance Tracking Task, "UTIAS Report 138, April 1969.
18. Innocenti M., "Interaction Between Display Dynamics and Handling Qualities in Manual Control Tasks", AFWAL-TM-84-215-FIGC, to be published.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

COMPUTATIONAL STUDY OF RAMJET COMBUSTOR FLOWFIELDS

Prepared by:	Dr. Kakkattukuzhy M. Isaac
Academic Rank:	Assistant Professor
Department and University:	Department of Mechanical and Aerospace Engineering, University of Missouri - Rolla
Research Location: Beam	Wright Patterson Air Force Base, Aero Propulsion Lab. Ramjet Technology Branch
USAF Research	Dr. A. S. Nejad
Date:	August 31, 1984
Contract No:	F49620-82-C-0035

COMPUTATIONAL STUDY OF RAMJET COMBUSTOR

FLOW FIELDS

by

Kakkattukuzhy M. Isaac

ABSTRACT

Numerical study of the typical ramjet/turbojet combustor flowfield has been carried out. The predictions show good agreement with the recently reported experimental data. The improvement in the prediction result from more realistic specification of the initial conditions and differences in the turbulence model used. Areas of deficiency in the computational method are identified and suggestions for further research for removing these deficiencies are offered.

#### ACKNOWLEDGEMENTS

The author wishes to thank the Air Force Systems command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing the opportunity to do research at the Aero Propulsion Lab, Wright Patterson Air Force Base, Dayton, Ohio. The facilities and help rendered by the Aero Propulsion Lab, in particular, the Ramjet Technology Branch is gratefully acknowledged.

The author wishes to thank Dr. A. S. Nejad for the keen interest shown and also for the helpful suggestions and fruitful discussions. He also would like to acknowledge the support, enthusiasm and assistance provided by Dr. R. R. Craig and Dr. F. D. Stull of the Ramjet Technology Branch.

## I. INTRODUCTION

Combustion chamber designs for turbojet and ramjet engines have evolved over the decades and have resulted in improved designs with higher levels of performance. The designers are still faced with several trial and error procedures due to the uncertainties in predictive capability. The reasons for this are many. The interplay of several physical phenomena such as turbulence, multiphase flow and chemical reaction make the flowfield very complex and, therefore, difficult to analyze. Very often the flowfield is three-dimensional in nature, which imposes severe limitations on computational capabilities in terms of storage and computation time. When a computer code is successfully developed, it is often found that there are few experimental data available to meaningfully prescribe the boundary conditions or to compare with the predictions.

Some encouraging developments have taken place in recent years which address some of the problems discussed above. Good progress has been made in the numerical modeling of the (Navier-Stokes) equations governing the recirculating flowfields typical of a ramjet/turbojet combustion chamber. Turbulence models capable of predicting the flowfields in a wide range of physical situations such as free shear flows (jets and wakes) wall-bounded flows (boundary-layers) and confined jets, have been developed. Chemical reaction models of varying degrees of sophistication are currently being used for combusting flows. These models range from a reactants-and-products equilibrium chemistry model to finite rate chemistry models with multi-step reactions involving several chemical species. Recent developments in non-intrusive experimental techniques have made it possible to make accurate measurements of mean velocity and turbulence even in highly turbulent, recirculating flow regions. Therefore, it is now possible to more accurately specify the boundary conditions for the solution of the equations; one is also

able to make meaningful comparisons between experiments and predictions without unduly worrying about experimental uncertainties.

Some combustor configurations that are currently being investigated are the dump combustor and the center bluff body combustor (Figures 1 and 2). The dump combustor appears to be particularly attractive due to its ability to decelerate the flow and create recirculating zones for flame holding and flame stabilization without significant pressure loss penalty. The present investigation was undertaken to evaluate the presently available predictive methods, particularly their performance with regard to the numerical methods and the turbulence model, in predicting confined, axisymmetric flowfields. The purpose was also to specify the boundary conditions as accurately as possible; previous predictive efforts were often plagued by improper specification of the inlet boundary conditions. The importance of properly specifying the boundary conditions has been emphasized in the past; however, the absence of reliable and sufficient data has precluded its implementation. A meaningful evaluation of the predictive scheme, nevertheless, depends heavily on how well the inlet conditions are specified. In what follows we have tried to address some of the problems discussed above.

## II. BACKGROUND

The numerical solution of the time-averaged, fully elliptic Navier-Stokes equations pose several difficulties because of the coupled, non-linear nature of the equations. The solution becomes particularly cumbersome when the flowfield considered is multidimensional, turbulent, compressible and chemically reacting. Turbulence introduces additional complexities because of what is known as the 'closure' problem. In compressible flows an additional equation (the energy equation) must be solved and the fluid density becomes a variable and it must be calculated from an equation of state. Chemical

reaction introduces additional equations for the transport of various chemical species.

The problems of turbulence modeling are still intriguing investigators in the field. One approach which has gained wide popularity is the  $k-\epsilon$  model<sup>1</sup> where  $k$  and  $\epsilon$  represent the turbulent kinetic energy and its dissipation rate, respectively. The  $k-\epsilon$  model treats turbulence by means of transport equations for these two variables. Several versions of the  $k-\epsilon$  model are in use; however, none of them gives uniformly good results for vastly varying flow conditions. Other approaches such as Reynolds stress models and pdf models can be found in the literature<sup>1-10</sup>. It appears that the  $k-\epsilon$  model offers more satisfactory results over a wide range of flow conditions in comparison to the other turbulence models which give good predictions in some cases and perform poorly in others.

The other aspect of the problem that we concern ourselves with, is the numerical algorithm for solving the set of non-linear, coupled, partial differential equations governing the flowfield. Solutions of the two-dimensional flow problems have been attempted in the past with vorticity and streamfunction as the dependent variables. A more widely used approach is due to Patankar and Spalding<sup>11</sup> who used the SIMPLE (Semi - Implicit - Pressure - Linked - Equations) algorithm for the solution of the equations in primitive variables. Their method has several features in common with the earlier work of Harlow and Welch<sup>12</sup>. Some of the deficiencies of these methods have been widely recognized and recently some alternate methods have been proposed<sup>13-15</sup> to overcome these deficiencies. Some of these methods are still in the development stage and, therefore their applicability is yet to be tested. The SIMPLE algorithm has been widely used for the solution of the Navier - Stokes equations governing parabolic as well as elliptic (recirculating) flows. Despite its shortcomings, the method has

successfully predicted a wide range of flowfields. Several investigators have used the method to predict the flowfield in combustion chambers of different geometries<sup>16-21</sup>.

### III. RESEARCH OBJECTIVE

The present investigation was undertaken with the following objectives:

- . Determine why the computer codes previously available at AFAPL (e.g. TEACH, STARPIC, STARRC) predict the flowfield quite different in nature from the experimentally observed ones in terms of center-line velocity decay, size of the recirculation zones and the radial velocity profiles.
- . Evaluate the  $k-\epsilon$  turbulence model in terms of its usefulness in predicting combustor flowfields.
- . Evaluate the numerical algorithm with respect to convergence and mesh specification.
- . Study the effect of upstream boundary conditions on the converged solution.
- . Suggest possible areas of improvements in terms of turbulence modeling and the numerics.
- . Suggests what data are critically needed to specify the inlet conditions so that the contribution of the uncertainty in the inlet conditions on the solution can be removed.

### IV. SCOPE

The work reported here is on the numerical investigation of the flowfield in a dump combustor using the time averaged Navier - Stokes equations for axisymmetric, recirculating flows. A two equation ( $k-\epsilon$ ) turbulence model was used for the calculations. The numerical implementation is based on the SIMPLE algorithm and a staggered grid approach as given in reference 22.



## V. GOVERNING EQUATIONS AND NUMERICAL MODELING

The equations governing the compressible, axisymmetric, combusting flow of a gas can be written as<sup>23</sup>

$$\frac{D\rho}{Dt} + \rho(\nabla \cdot \bar{V}) = 0 \quad (1)$$

$$\rho \frac{D\bar{V}}{Dt} = \rho \bar{F} - \nabla p + \nabla \cdot \bar{\bar{T}} \quad (2)$$

$$\rho \frac{Dh}{Dt} = -\nabla \cdot \bar{J}_h - \rho(\nabla \cdot \bar{V}) + \nabla \bar{V} \cdot \bar{\bar{T}} \quad (3)$$

$$\rho \frac{Dm_j}{Dt} = -\nabla \cdot \bar{J}_j + R_j \quad (4)$$

In these equations, the symbols have the following meaning.

- $p$  - pressure
- $\rho$  - density
- $\bar{V}$  - velocity
- $\bar{F}$  - body force per unit mass
- $\bar{\bar{T}}$  - turbulent stress tensor
- $h$  - enthalpy
- $\bar{J}_h$  - turbulent flux vector for enthalpy
- $m_j$  - time-mean chemical species mass fraction
- $\bar{J}_j$  - turbulent flux vector for species,  $j$ .
- $R_j$  - mass rate of creation of species,  $j$ .

In addition we have the following thermodynamics relations.

Equation of state:

$$p = \rho RT \quad (5)$$

Enthalpy:

$$h = C_p T + \sum_j H_j m_j \quad (6)$$

where  $C_p$  is the constant pressure specific heat and  $H_j$ , the enthalpy of formation of species,  $j$ .

Rate Equation:

$$R_j = -F_j \exp(-E_j/RT) \quad (7)$$

with the symbols having the following meaning:

$F_j$  - frequency factor

$E_j$  - activation energy

$R$  - universal gas constant.

Equations 1 to 4 are further simplified by assuming isotropy and constant values of turbulent Prandtl number and Schmidt number.

In order to implement the numerical scheme efficiently equations 1 to 4 are cast in the general form:

$$\frac{1}{r} \left[ \frac{\partial}{\partial x} (\rho u r \phi) + \frac{\partial}{\partial r} (\rho v r \phi) - \frac{\partial}{\partial x} (r \Gamma_\phi \frac{\partial \phi}{\partial x}) - \frac{\partial}{\partial r} (r \Gamma_\phi \frac{\partial \phi}{\partial r}) \right] = S_\phi \quad (8)$$

where  $\phi = 1$  for continuity equation,  $\phi = V$  for the three components of momentum equation,  $\phi = h$  for energy equation and  $\phi = m_j$  for each of the transport equations for the species.

In the  $k-\epsilon$  model, turbulence closure is achieved by having two more transport equations for  $k$  and  $\epsilon$  which are also in the form given in equation 8. The source term,  $S_\phi$  on the RHS has a different form for each of the variable. These can be found in references 16-21. For the present calculations, the source term  $S_\phi$  is zero for the energy equation.

The partial differential equations for continuity, momentum, energy and species represented by the general form given by equation 8 are solved by the SIMPLE algorithm of reference 22. Details of the solution procedure may be found in several other references also<sup>16-21</sup>. The discretization equations are obtained by integrating the partial differential equations over the cell control volumes formed by the grid lines in the solution domain. A lucid explanation of its advantages over other differencing methods is given in reference 22. Suffice to state that the procedure ensures conservation of

quantities such as mass, momentum and energy over any group of control volumes and therefore, over the whole computational domain. A staggered grid approach and a central/upwind differencing scheme are used in applying the equations; the advantages of these can be found in reference 22. A line-relaxation procedure is used for the solution of the equations. Convergence is improved by reliance on under-relaxation factors which control the variation of the dependent variables from iteration to iteration.

The energy equation warrants particular mention. The compressible form of the energy equation is used for the present calculations and, therefore, density is treated as a variable. It is calculated from the equation of state:

$$p = \rho RT$$

Patankar<sup>22</sup> has given discretization equations for variable density flows; however, the present calculations do not use Patankar's method. Instead, convergence is achieved by means of restart tapes which take the solution from incompressible flows to compressible subsonic flows. The initial guess for a new case is obtained from the converged solution of a similar case obtained before. The coding was carried out to handle various types of boundary conditions for the energy equation. Adiabatic wall, constant temperature wall and constant heat flux wall are the presently available options. Solutions with temperature differences of upto 100C between the wall and the gas at the inlet has been obtained without encountering convergence problems. The computer code has been written keeping in mind that the future inclusion of chemical kinetics for combusting flows should be able to be carried out without difficulty.

## VI. RESULTS AND DISCUSSION;

The results of the study are plotted in Figures 3-5. The plots show the

center-line velocity decay, radial velocity profiles for five axial stations and the variation of the turbulent kinetic energy along the center-line.

In Figure 3, the present predictions using the computer code ECFLOW and the predictions using an existing code (STARPIC) are compared with the experimental data of reference 24. It is instructive to examine the differences between the present predictions and the earlier ones. The source of these errors may be traced to the differences in specifying the inlet boundary conditions for the  $k$  and the  $\epsilon$  equations, the differences in the constants used in the  $k$ - $\epsilon$  model and some coding errors in the computer code STARPIC. (e.g. See expression for the diffusion coefficient 'GAMN' in STARPIC).

The present calculations used the data of Craig, Nejad et al<sup>24</sup> for specifying the turbulent kinetic energy ( $k$ ) at the inlet; since only the axial component of the turbulent intensity was measured by these investigators, it was necessary to assume isotropy for the calculation of  $k$  from the axial turbulent intensity. The energy dissipation rate  $\epsilon$  was specified in terms of the macro length scale,  $l$  and a constant  $C_D$ . It is given by:

$$\epsilon = C_D k^{3/2} / l.$$

The  $k$ - $\epsilon$  model has been well tested for free-shear flows such as jets and wakes. It is a usual practice to specify the macro eddy-length scale in such flows to be of the same order as the jet/wake half width. There is no corresponding obvious way of specifying the length scale for internal flows. In confined flows it is reasonable to expect the length scale to be smaller because of the confining nature of the walls. Therefore, the value of  $l$  was chosen to be smaller than that suggested by Leschiner and Rodi who used the expression:

$$l = 0.56 \delta$$

where  $\delta$  is the thickness of the shear layer. For the present calculations  $l$

was chosen to be equal to one tenth of the inlet diameter,  $d$ .

$$l = 0.1 d$$

This modification results in considerable improvement in the agreement between prediction and data; there is no other explanation for using this value for  $l$ .

It is seen from Figure 3 that the axial velocity is underpredicted in the entrance region ( See Figure 1 for a schematic of the different regions of the flowfield.) of the combustor. Since the velocity profile at the entrance is a top-hat profile representative of the exit flow from convergent nozzles, it is reasonable to assume the existence of a sizable potential core at the combustor entrance. Jones and Launder<sup>4</sup> have reported that the turbulence model of the form given in reference 1 is inadequate in predicting flowfields in regions of low turbulent Reynolds number given by:

$$R_t = \rho k^2 / \mu \epsilon$$

They have used additional terms in the transport equations for  $k$  and  $\epsilon$  and also made the constants  $C_1$  and  $C_2$  functions of  $R_t$ . The reason for the overprediction of the axial velocity in the entrance region could probably be attributed to the factors cited by Jones and Launder<sup>4</sup>. From the radial velocity profiles of Figures 4 it may be observed that there exists, near the axis, a region of flat velocity profile. The extent of this region gradually diminishes indicating that it represents the potential core. It may also be observed that the major differences between the prediction and the data are confined to this region of low turbulence intensities as well as the region of reverse flow.

Differences between prediction and measurements can also be seen in the reverse flow region (See the radial profiles of Figure 4; the reverse flow region is represented by the negative values of the axial velocity). For the axial station of  $X/D = 0.53$  the data show considerable scatter near the wall.

Since the position of the reattachment point could be estimated by observing where the radial profiles stop showing negative values of axial velocity near the wall, the size of the recirculation zone can be easily estimated. Thus, it is easily seen that computations under-predict the size of the recirculations zone; similar trend has also been reported by Krishnamoorthy and Park<sup>21</sup> who attributed the differences to the streamline curvature in the recirculation region. Modifications to the  $k-\epsilon$  model to account for streamline curvature has been suggested by Leschiner and Rodi<sup>25</sup>. These have been incorporated in reference 21 for predicting confined flow behind axisymmetric bluff bodies.

Figure 5 shows the variation of the turbulent kinetic energy ( $k$ ) along the combustor center-line. It is important to note that Craig, Nejad et al<sup>24</sup> measured only the axial turbulence intensity; therefore, the experimental values of turbulent kinetic energy shown in Figure 5 were calculated assuming isotropy. The data are too few to fit a curve; therefore, a tentative variation is shown by the dotted line. The solid curve represents the predicted values of the turbulent kinetic energy. The data and the predictions differ quite substantially in their absolute values; the reasons for this will become more clear when detailed measurements of the turbulence quantities become available. The curves, nevertheless, have similar trends. Both the curves show three distinct regions. The first is the initial region where the turbulent energy decays; this probably is the potential core and the flow in this region is likely to be isotropic. The next region is characterized by a sharp rise in the value of  $k$ , suggesting, may be, that the potential core has ended and the mixing layer has reached the center-line. Further downstream,  $k$  starts decaying indicating that fully developed turbulent pipe flow is getting established.

## VII RECOMMENDATIONS:

To conclude, we summarize the major findings of the present study and make a few recommendations for further work.

. The present work clearly indicates the significant influence of the boundary conditions on the predictions. It, therefore, underscores the importance of conducting carefully designed experiments to determine the initial conditions for validating the computer codes. A large data base is also needed to empirically establish the combustor inlet conditions in situations where it is not practical to make measurements.

. The originally proposed constants in the  $k-\epsilon$  turbulence model are not adequate for predicting recirculating flow in internal geometries. The present study shows that predictions can be improved by adjusting the constants.

. The turbulence model performs poorly in the potential core region at the entrance. This subsequently has a bearing on the solution everywhere. The turbulence model also must be modified for predicting flows in regions where the streamlines have large curvature such as in recirculating flow regions. I propose that modifications similar to the ones suggested by Jones and Launder be used for the potential core region.

. There is considerable disagreement between prediction and experiments in the velocity profiles in the recirculating flow region. The computations also underpredict the size of the recirculating region. I propose that further research be carried out using a curvature dependent damping factor to account for the reduction in the turbulent kinetic energy due to curvature of the streamlines.

#### REFERENCES

1. Launder, B. E. and D. B. Spalding, Mathematical Models of Turbulence (Academic Press, London, 1972).
2. Bradshaw, P. D. H. Ferriss, and N. P. Atwell, "Calculation of Boundary Layer Development Using the Turbulent Energy Equation," J. Fluid Mech., 28, 1967, pp. 593-616.
3. Hanjalic, K., Two-dimensional Asymmetric Turbulent Flow in Ducts, Ph.D. Thesis, University of London (1970).
4. Jones, W. P. and B. E. Launder, "The Prediction of Laminarization with Two-Equation Model of Turbulence," Intl. J. Heat and Mass Transfer, 15, 1972, pp. 301-304.
5. Lumley, T. L., and B. J. Khajeh Nouri, "Computational Modeling of Turbulent Transport," Adv. Geophys., Vol. 18A, 1974, pp. 169-192.
6. Wilcox, D. C. and M. W. Rubesin, "Progress in Turbulence Modeling for Complex Flow Fields Including the Effect of Compressibility," NASA TP-1517, 1980.
7. Gibson, M. M. and W. Rodi, "A Reynolds-stress Closure Model of Turbulence Applied to the Calculation of a Highly Curved Mixing Layer," Journal of Fluid Mechanics, Vol. 103, 1981.
8. Chien, J. Y., "Predictions of Channel Boundary-Layer Flows with a Low-Reynolds-Number Turbulence Model," AIAA J., 20, 1982, PP. 33-38.
9. Coakley, T. J., "Turbulence Modeling Methods of the Compressible Navier-Stokes Equations," AIAA Paper 83-1963, 1983.
10. Bywater, R. J., "Velocity Space Description of Certain Turbulent Free Shear Flow Characteristics," AIAA Journal, Vol. 20, No. 6, June 1982.
11. Patankar, S. V. and D. B. Spalding, Heat and Mass Transfer in Boundary Layers, 2nd ed., Intertext, London.
12. Harlow, F. H. and J. E. Welch, "Numerical Calculation of Time-dependent Viscous Incompressible Flow," Phys. Fluids, Vol. 8, 1965, pp. 2182-2189.
13. Raithby, G. D., "Skew Upstream Differencing Schemes for Problems involving Fluid Flow," Computer Methods in Applied Mechanics and Engineering 9, 1976.
14. Leonard, B. P., "Stable and Accurate Convective Modeling Procedure Based on Quadratic Upstream Interpolation," Computer Methods in Applied Mechanics and Engineering, Vol. 19, 1979.
15. Vanka, S. P. and G. K. Leaf, "Efficient Finite Difference Calculation Procedure for Multi-Dimensional Fluid Flows," AIAA-84-1244, AIAA/SAE/ASME 20th Joint Propulsion Conference, June 1984.



#### REFERENCES (Continued)

16. Gosman, A. D. and F. J. K. Ideraih, "TEACH-2E: A General Computer Program for Two-Dimensional Turbulent Recirculating Flows," Report, Dept. of Mech. Eng., Imperial College, 1976.
17. Lilley, D. G. and D. L. Rhode, "A Computer Code for Swirling Turbulency Axisymmetric Recirculating Flows in Practical Isothermal Combustor Geometries," NASA-CR-3442, 1982.
18. Boray, R. S. and C. Chang, "Flow Studies of Dump Combustors," Fifth International Symposium on Airbreathing Engines," February 16-21, 1981, Bangalore, India.
19. Sindir, M. M. and P. T. Harsha, "Assessment of Turbulence Models for Scramjet Flowfields," SAI Rept. No. 82-038-CHA, 1982.
20. Harsh, W. H., "Numerical Modeling of Ramjet Combustors," AFWAL-TR-82-2113, Feb. 1983.
21. Krishnamurthy, L. and S. O. Park, "Streamline Curvature Effects in Confined Isothermal Recirculating Flowfields Behind an Axisymmetric Bluff Body: Numerical Calculations with the k- Turbulence Model," Fourth Symposium on Turbulent Shear Flows, Karlsruhe, F. R. Germany, September 1983.
22. Patankar, S. V., Numerical Heat Transfer and Fluid Flow, Hemisphere, Washington, DC, 1980.
23. Bird, R. B., W. E. Stewart and E. N. Lightfoot, Transport Phenomena, John Wiley & Sons, 1960.
24. Craig, R. R., A. S. Nejad, E. Y. Hahn, and K. G. Schwertzkopf, "A General Approach for Obtaining Unbiased LDV Data in Highly Turbulent Non-reacting and Reacting Flows," AIAA 22nd Aerospace Sciences Meeting, AIAA-84-0366.
25. Leschziner, M. A. and W. Rodi, "Calculation of Annular and Twin Parallel Jets Using Various Discretization Schemes and Turbulence-Model Variations," J. Fluids Engng., 103, A352, 1981.

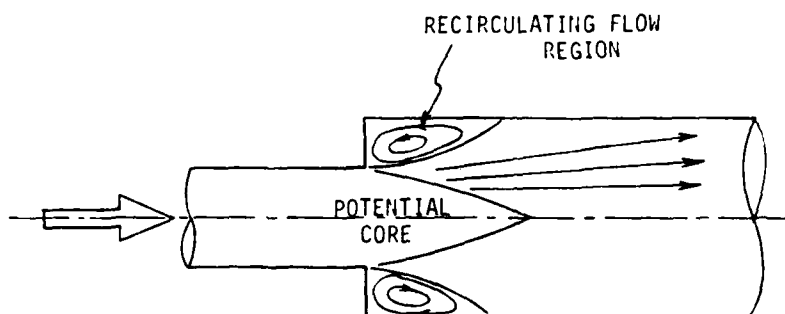


FIGURE 1. SCHEATIC OF DUMP COMBUSTOR

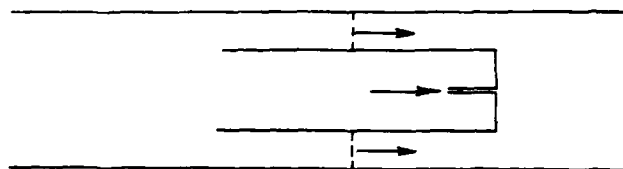


FIGURE 2. SCHEMATIC OF CENTER BLUFF BODY COMBUSTOR

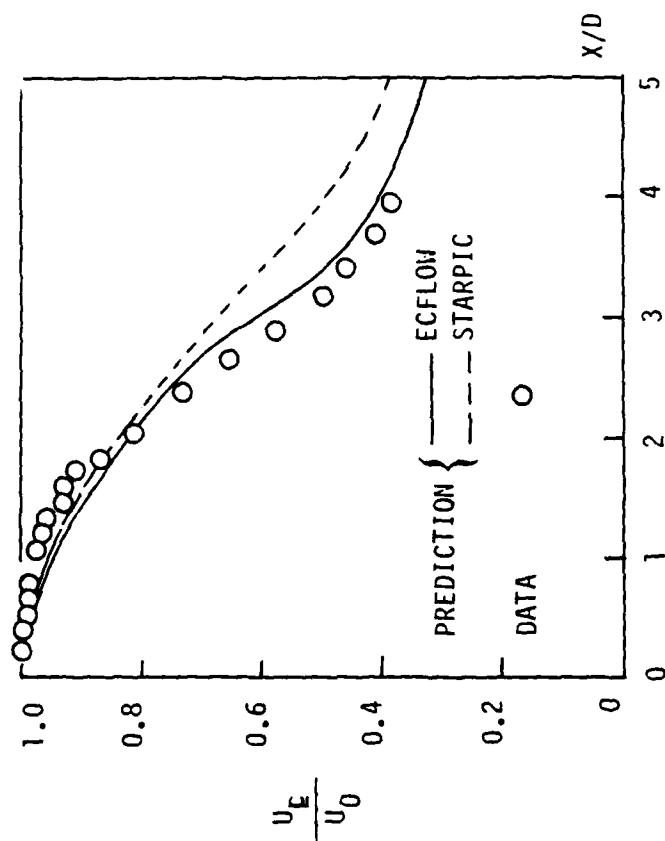


FIGURE 3 - Variation of axial velocity along combustor center-line-comparison between prediction and experiments.

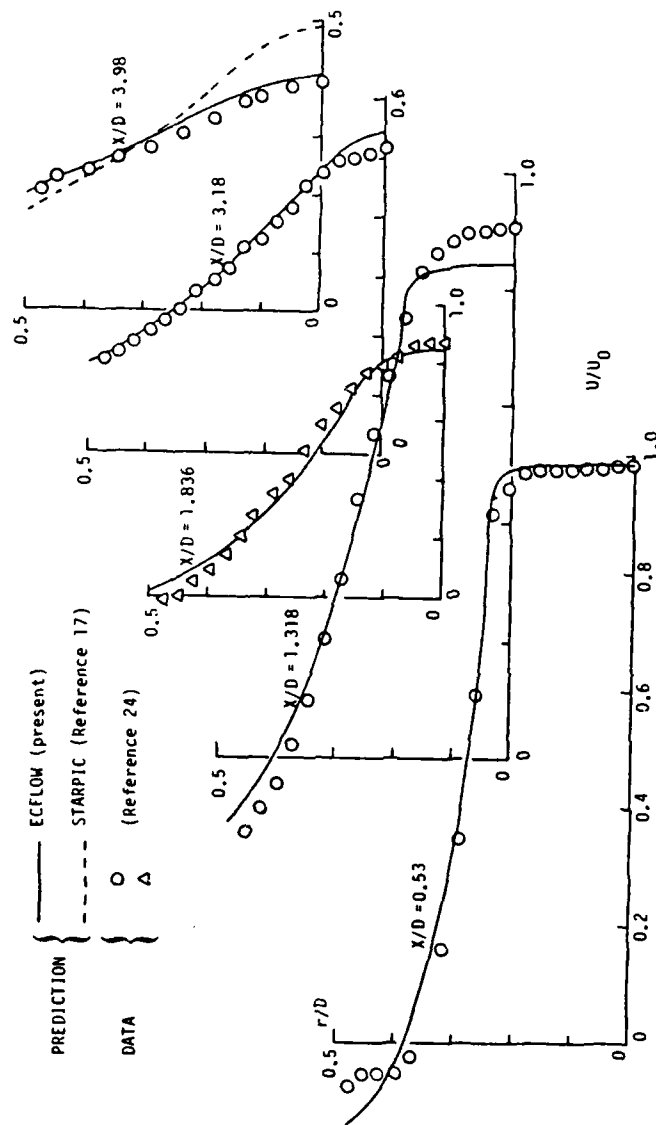


FIGURE 4 - Radial velocity profiles at different axial stations - comparison of predictions with experiments.

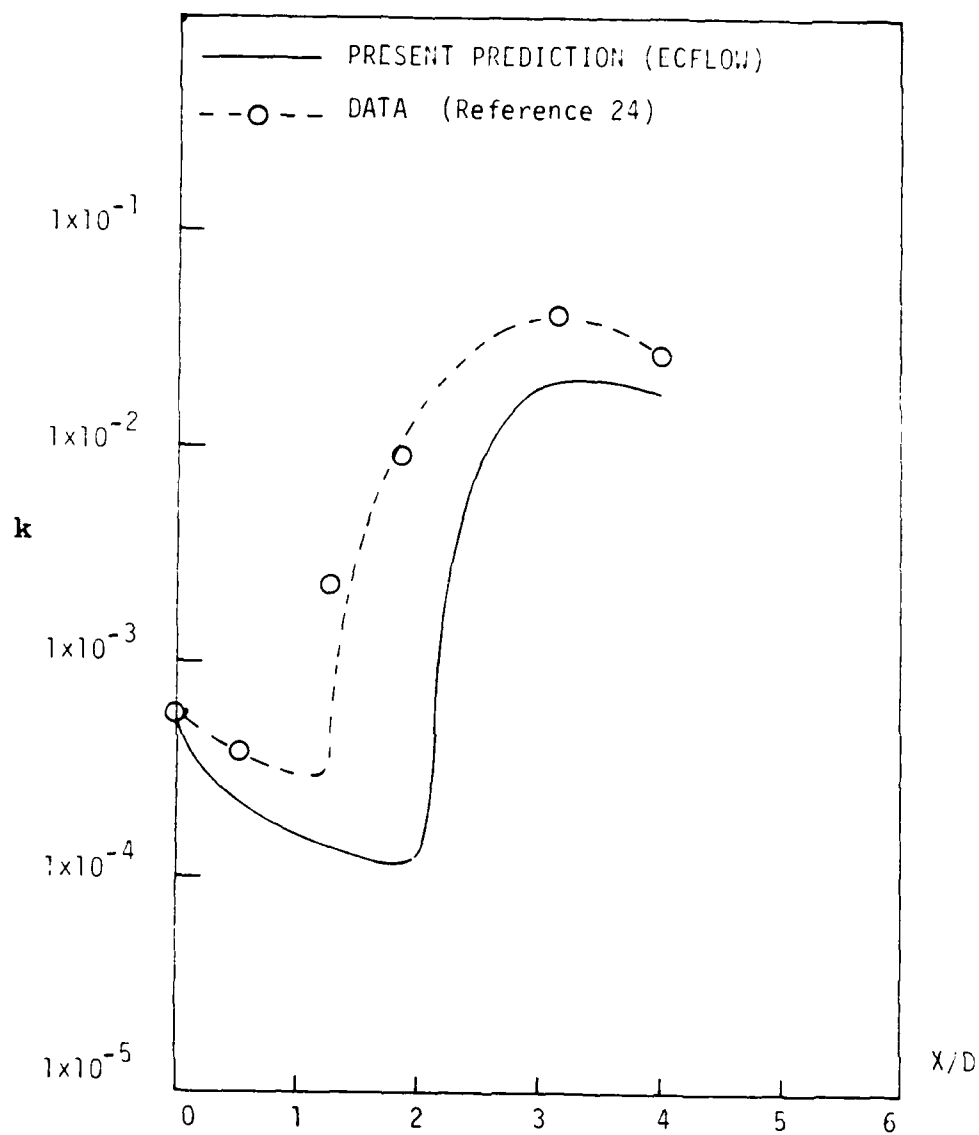


FIGURE 5 - Variation of turbulent kinetic energy along combustor center line - comparison between prediction and experiments.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING  
EDUCATION

FINAL REPORT

NUMERICAL CHARACTERIZATION OF MICROSTRIP  
DISCONTINUITIES ON THICK SUBSTRATES

Prepared by: Dr. Robert W. Jackson

Academic Rank: Assistant Professor

Department and Department of Electrical and Computer  
University: Engineering, University of Massachusetts

Research Location: Rome Air Development Command  
Electromagnetic Sciences Division/E.E.A.  
Hanscom Air Force Base

USAF Research Dr. John K. Schindler

Date: September 20, 1984

Contract No.: F49620-82-C-0035

Numerical Characterization of Microstrip  
Discontinuities on Thick Substrates

by

Robert W. Jackson

Abstract

Microstrip discontinuities have heretofore been characterized by models which do not include surface wave or radiation effects. In this work, the microstrip step discontinuity is investigated by numerically solving an integral equation which was derived using the grounded dielectric slab Green's function. This technique includes the surface wave and space wave effects which are expected to be important on thick substrates. Using a moment method technique the current is expanded in terms of longitudinal sinusoidal and piecewise sinusoidal modes. S parameters are obtained. Results are promising but indicate that transverse currents may be important.

## I. Introduction

This research investigates the characterization of microstrip discontinuities on electrically thick substrates. Interest in this work stems from the uncertain validity of discontinuity models used in standard computer aided design routines such as SUPERCOMPACT. The use of computer aided design is a necessity for efficient low cost design of microwave and millimeter wave integrated circuits. Present discontinuity models appear to be valid on electrically thin substrates. At lower frequencies (X band), an electrically thin substrate is still physically thick enough to process. At millimeter wave frequencies, electrically thin substrates made of materials such as Gallium Arsenide would be too thin to process without breaking or warping. Therefore computer design routines will be needed for discontinuities on thick substrates.

Models for microstrip discontinuities are currently generated from numerical analysis based upon quasistatic assumptions [1], [2]. This type of analysis does not include surface wave or radiation effects. The "Planar Waveguide" model [3] has been developed to include the effect of dispersion and higher order models, but still does not include loss to surface waves or radiation. Spectral Domain techniques [4] have also been used but since these techniques have been applied to a discontinuity enclosed in a perfectly conducting box, they do not accurately account for surface wave and radiation effects.

The technique used in this research involves developing an integral equation for the current on a microstrip step discontinuity. The formulation uses the grounded dielectric slab Green's function which is valid on open structures and which includes surface wave and radiation



effects. This method has been used by Pozar [5] to analyze the open ended microstrip. The integral equation formed is solved numerically via the method of moments.

The following section reviews the objectives of this research. Section III formulates the integral equation and the general method of solution. Section IV discusses the various choices of expansion functions and the results. Section V gives conclusions and recommendations for future work.

## II. Objectives

The general objective of this research program was to investigate computer modeling of microstrip discontinuities on electrically thick substrates. The specific effects which are included in this type of analysis but not in previous work are surface wave losses, radiation losses and a more rigorous analysis of dispersive effects.

The specific objectives for this summer program are:

- (1) Modify the moment method technique used by Pozar on the open ended microstrip so that it can be applied to microstrip step discontinuity.
- (2) Set up a computer routine which will efficiently solve the set of equations generated in objective (1). This is non-trivial due the fact that to determine matrix elements for this set of equations one must do a double integration of a complex integrand which varies rapidly (Sommerfeld integral). If not efficiently designed the software routine could use a large amount of computer time.
- (3) Compare the generated results to known quasistatic results.

- (4) Investigate the application of this technique to other microstrip discontinuities.

### III. Theoretical Framework

In this section the integral equation for the step discontinuity is formulated and the moment method approach to the solution is discussed. The problem is as follows. A perfectly conducting microstrip of width  $w$  lies on a grounded dielectric slab of thickness,  $d$ . At  $x=0$ , the width increases to  $w_+$  (see figure 1). We wish to find the scattering parameters for this junction.

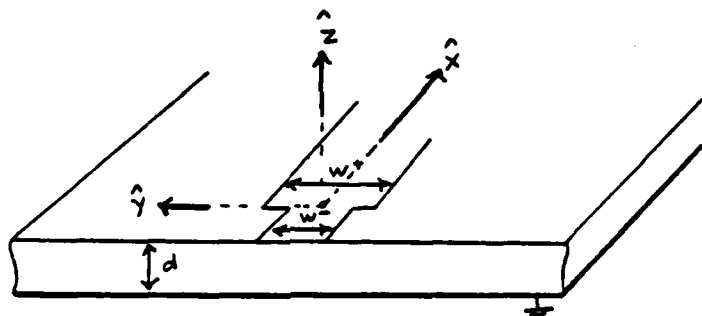


figure 1.

To derive the integral equation we start with the expression for the  $x$  directed electric field on the dielectric/air interface due to an arbitrary  $x$  directed surface current,  $J_x$ , also on the dielectric/air interface.

$$E_x(x,y,d) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(x,y,x_0,y_0) J_x(x_0,y_0,d) dx_0 dy_0 \quad (1)$$

where  $G(x,y,x_0,y_0)$  is the grounded dielectric slab Green's function,

$$G(x, y, x_0, y_0) = \frac{-jZ_0}{4\pi^2 k_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(k_x, k_y) e^{jk_x(x-x_0)} e^{jk_y(y-y_0)} dk_x dk_y, \quad (2)$$

where

$$Q(k_x, k_y) = \frac{(\epsilon_r k_0^2 - k_x^2) k_2 \cos k_1 d + j k_1 (k_0^2 - k_x^2) \sin k_1 d}{T_e T_m} \sin k_1 d, \quad (3)$$

$$T_e = k_1 \cos k_1 d + j k_2 \sin k_1 d,$$

$$T_m = \epsilon_r k_2 \cos k_2 d + j k_1 \sin k_1 d,$$

$$k_1^2 = \epsilon_r k_0^2 - \beta^2, \quad (\text{Im} k_1 < 0)$$

$$k_2^2 = k_0^2 - \beta^2, \quad (\text{Im} k_2 < 0)$$

$$\beta^2 = k_x^2 + k_y^2,$$

$$Z_0 = \sqrt{\mu_0 / \epsilon_0}.$$

The fields generated by equation (1) satisfy all boundary conditions at the dielectric interface and at the ground plane. Surface waves are included exactly and come about as a result of the zeros in the denominator of (3). The only boundary condition which is not satisfied occurs if conductors are present at the dielectric/air interface.  $J_x$  must then be constrained such that  $E_x$  is zero on the conductors. For the step discontinuity,  $J_x$  is split into two parts, an impressed sinusoidal current  $J_x^i$  which excites the junction and the excited current  $J_x^e$  which includes the reflected and transmitted waves. By forcing  $E_x$  in equation (1) to be zero on the microstrip we arrive at the following integral equation,

$$\int_{-w}^w \int_{-w}^w G(x, x_0, y, y_0) [J_x^e(x_0, y_0, d) + J_x^i(x_0, y_0, d)] dx_0 dy_0$$

$$= 0 \text{ for } \begin{cases} |y| < w^-/2, x < 0 \\ |y| < w^+/2, x > 0 \end{cases} \quad (4)$$

To implement the moment method solution, the unknown current,  $J_x^e$ , is expanded to a sum of known functions multiplied by unknown constants. Sufficient equations to uniquely determine these constants are generated by taking weighted averages (moments) of equation (4) over all  $x, y$  space. We expand the currents in the following way,

$$J_x^e = \sum_{n=1}^N I_n^- f_n^-(x-x_n, y) + \sum_{m=1}^M I_m^+ f_m^+(x-x_m^+, y)$$

$$+ I^{CS-} [f^{C-}(x, y) + j f^{S-}(x, y)] + I^{CS+} [f^{C+}(x, y) - j f^{S+}(x, y)] \quad (5)$$

$$J_x^i = f^{C-}(x, y) - j f^{S-}(x, y) \text{ or } f^{C+}(x, y) + j f^{S+}(x, y) \quad (6)$$

where  $\{I_n^-, I_m^+, I^{CS-}, I^{CS+}\}$  are unknown and where we have chosen

$$f_n^i(x-x_n, y) = \frac{\sin[h^i(\alpha^i - |x-x_n|)]}{w^i \sin(h^i \frac{\alpha^i}{2})} \text{ for } \begin{cases} |x-x_n| < \alpha^i \\ |y| < w^i/2 \end{cases} \quad (7)$$

$$f^{Ci}(x, y) = \cos(h^i x) \text{ for } \begin{cases} 0 < [(.5\pi/h^i) + x] < T^i \\ |y| < w^i/2 \end{cases} \quad (8)$$

$$f^{Si}(x, y) = \sin(h^i x) \text{ for } \begin{cases} 0 < [-x] < T^i \\ |y| < w^i/2 \end{cases} \quad (9)$$

$$i = \begin{cases} -, x < 0 \\ +, x > 0 \end{cases}$$

In the near vicinity of the junction the piecewise sinusoidal modes,  $f_n^{\pm}$ , centered around  $x_n$  are used to approximate the current. Farther away, finite length sinusoids are used. The constants  $\{a^-, a^+\}$  are the half lengths of the piecewise sinusoids and  $\{h^-, h^+\}$  are the propagation constants for the  $z < 0$  and  $z > 0$  microstrips.  $\{T^-, T^+\}$  are the lengths the finite sinusoids—usually 3 or 4 wavelengths. The weighting functions are all chosen to be the piecewise sinusoids,  $\{f_n^+, f_n^-\}$ . Multiplying equation (4) by the weighting functions and integrating over  $x$  and  $y$  yields the following set of simultaneous equations,

$$\sum_{n=1}^N Z_{jn}^- I_n^- + \sum_{m=1}^M Z_{jm}^+ I_m^+ + Z_j^{cs-} I_j^{cs-} - Z_j^{cs+} I_j^{cs+} = V_j \quad (10)$$

$$j = 1, 2, \dots, N+M+2$$

where

$$\begin{Bmatrix} Z_{jn}^- \\ Z_{jn}^+ \\ Z_j^{cs-} \\ Z_j^{cs+} \end{Bmatrix} = - \int_{-a}^a \int_{-a}^a dx dy \int_{-a}^a dx_0 dy_0 f_j^k(x-x_j^k, y) G(x, x_0, y, y_0) \begin{Bmatrix} f_n^-(x-x_n, y) \\ f_n^+(x-x_n^+, y) \\ f^{cs-} \\ f^{cs+} \end{Bmatrix} \quad (11)$$

$$k = \{-, +\}$$

$$V_j = \int_{-a}^a \int_{-a}^a dx dy \int_{-a}^a dx_0 dy_0 f_j^k(x-x_j^k, y) G(x, x_0, y, y_0) J_x^1(x, y) \quad (12)$$

where the  $i$  superscript indicates whether the weighting function is a  $z < 0$  or  $z > 0$  piecewise sinusoid (the lengths and widths differ). The  $Z$  matrix and  $V$  coefficients can be simplified by doing the spatial integrations analytically. For a typical case

$$Z_{jn}^{-} = \frac{jZ_0}{4\pi^2 k_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(k_x, k_y) F_j^{+}(k_x, k_y) F_n^{-*}(k_x, k_y) dk_x dk_y \quad (13)$$

where

$$F_j^{-}(k_x, k_y) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy f_j^{-}(x-x_j, y) e^{jk_x x} e^{jk_y y},$$

the fourier transform of  $f_j^{-}$ , is done analytically. The fourier transforming not only reduces the total number of integrations but makes it possible to easily shift current elements in space with a simple multiplication by  $\exp(jk_x \Delta x)$ .

In summary, matrix elements similar to (13) are evaluated numerically then used to solve matrix equation (10) for  $I^{cs-}$  and  $I^{cs+}$ , the amplitudes of the excited sinusoidal waves.

Two necessary inputs are the propagation constants for the microstrip line on either side of the junction. A dispersion relation for these constants can be derived in the following manner. Assume the current on an infinite microstrip of width,  $w$  is

$$J_x(x, y, d) = \begin{cases} e^{-jhx} & , \quad |y| < w/2 \\ 0 & , \quad |y| > w/2 \end{cases}$$

and substitute into equation (1). The resulting field has the form

$$E_x(x, y, d) = E(y) e^{-jhx} \quad (14)$$

the amplitude  $E(y)$  must be zero on the microstrip. To enforce this multiply  $E(y)$  by the testing (weighting) function,

$$W(y) = \begin{cases} 1 & , \quad |y| < w/2 \\ 0 & , \quad |y| > w/2 \end{cases}$$

integrate over  $y$ , and set to zero. The resulting dispersion relation is,

$$\frac{-jZ_0}{4\pi^2 k_0} \int_{-\infty}^{\infty} dk_y Q(k_x = -h, k_y) \left[ \frac{\sin[k_y w/2]}{k_y} \right]^2 = 0 \quad (15)$$

The integral is evaluated numerically while varying  $h$  until a zero occurs. Surface wave effects and all fringing fields are included.

Expansion currents were chosen as simply as possible. Figure 2 shows an example. Currents were constant in the

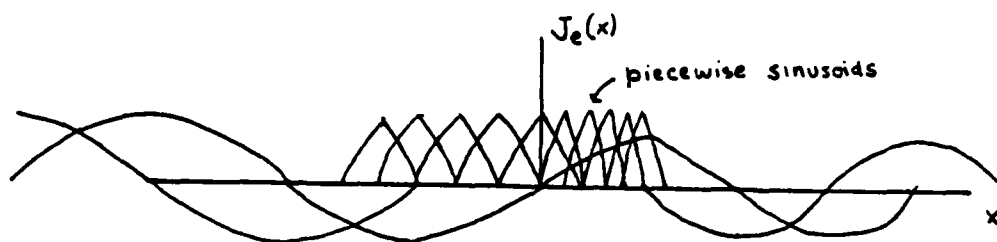


Figure 2. A set of expansion currents

transverse direction since transverse current variations have been found to be unimportant in moment method dipole antenna results [5]. For simplicity all currents were taken to be  $x$  directed.

#### IV. Numerical Considerations

The formulations used in this research, the spectral approach and the moment method, are very rigorous but there are well known drawbacks. We briefly discuss the problems most relevant to this work.

Calculation of the impedance elements in equation (13) involves an integration over an infinite area in  $\{k_x, k_y\}$  space. We convert this to an  $\alpha, \beta$  integration where

$$k_x = \beta \cos \alpha$$

$$k_y = \beta \sin \alpha$$

Due to certain symmetries, the integration can be reduced to an integration over  $0 < \beta < \infty$  and  $0 < \alpha < \pi/2$ . A singularity occurs in the  $\beta$  integration due to a surface wave pole. Evaluation of this singularity is discussed elsewhere [6]. The integrand varies rapidly enough that as many as 800 integrand evaluations were necessary per  $\beta$  point near the upper  $\beta$  integration limit. The integrals converged sufficiently for the integration to be truncated at an upper  $\beta$  limit of  $100k_0$  and less on thicker substrates. Choosing the sinusoidal wave trains to be an integral number of half wavelengths long helped speed the convergence by a factor  $\beta^{-2}$  in the integrand. Another large time saving occurs if, for each  $\{\alpha, \beta\}$  point in the integration, all impedance matrix terms are evaluated at the same time. This takes advantage of the fact that several factors in the integrand are common to all impedance matrix elements.

The other numerical consideration in this work involves the moment method. The approximate current hopefully converges to the real current if enough expansion modes are added. The results in this work were considered promising if they were insensitive to adding additional modes or to an



decrease in mode size. Since moment method results can converge to an incorrect result, results are compared to other accepted results in common regions of applicability. In this case results for thin substrates were compared to quasistatic results.

#### IV. Results

Results were generated for two types of discontinuities, a microstrip gap and the microstrip step. The microstrip gap was chosen since it is somewhat simpler than a step and since the results better agree with quasistatic results in the thin substrates limit. Results on step discontinuity were not conclusive and will be discussed later.

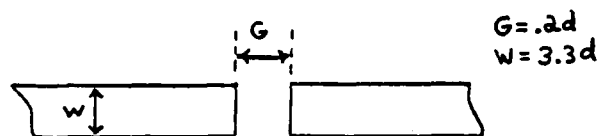
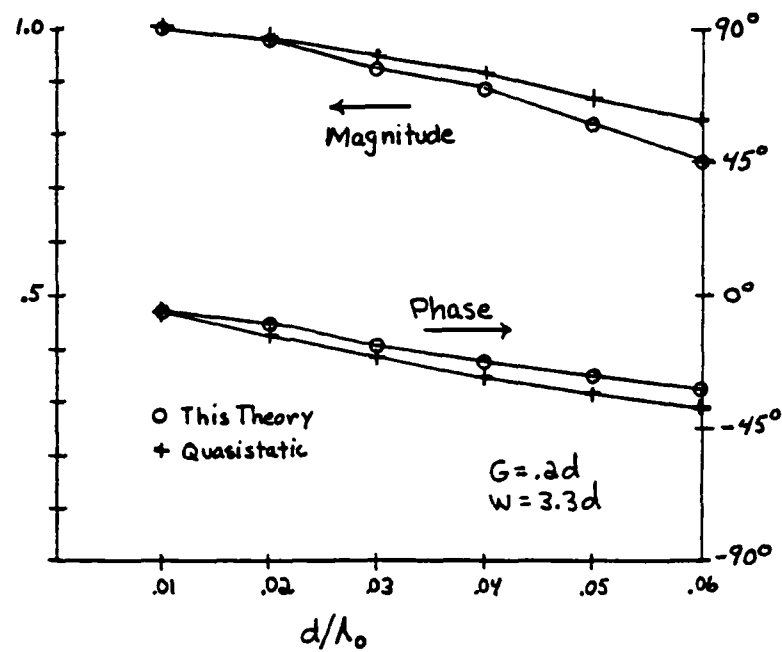


Figure 3. Gap Junction

Figure 3 is the representative structure for which  $S_{11}$  and  $S_{21}$  are calculated and plotted versus electrical thickness (frequency) in figure 4.  $S_{22}$  and  $S_{12}$  are almost exactly the same as  $S_{11}$  and  $S_{21}$ , respectively. Quasistatic results from the gap analysis routine in SUPERCOMPACT are given for comparison. Notice that the magnitude of  $S_{11}$  and the angles of  $S_{11}$  and  $S_{21}$  coincide almost exactly with the quasistatic results for thin substrates. The magnitude of  $S_{21}$  does not closely match the quasistatic

S11



S21

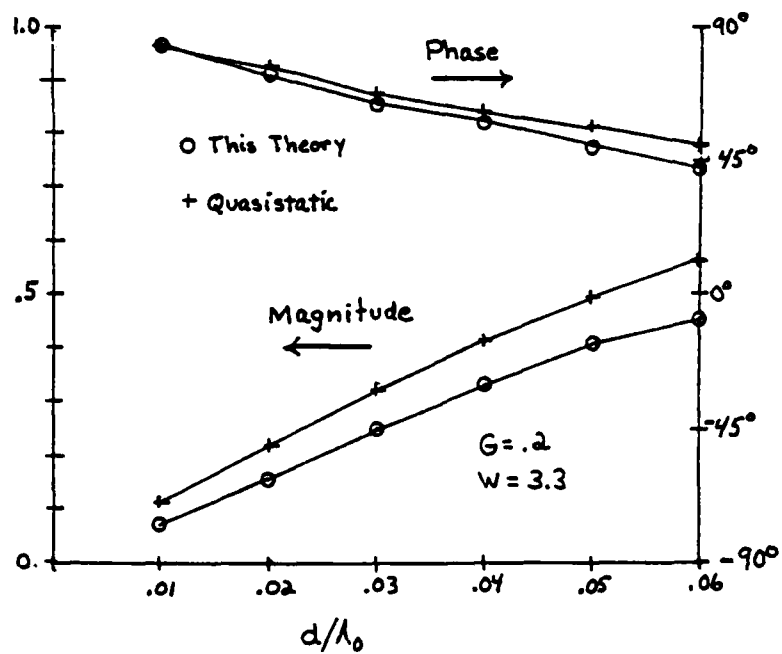


Figure 4 S parameters for Microstrip Gap

results but does follow the general trend. In Table 1 we list the quantity  $|S_{21}|^2 + |S_{11}|^2$  vs. electrical thickness. This quantity is unity for the lossless quasistatic result. The table shows increased loss to radiation and surface waves for electrically thicker substrates.

$d/\lambda_0$	$ S_{11} ^2 +  S_{21} ^2$
.01	.99
.02	.97
.03	.93
.04	.88
.05	.83
.06	.78

Table 1. Measure of Loss for Gap Discontinuity

The step discontinuity was investigated for the junction between microstriplines of approximately 100 ohms and 50 ohms on  $\epsilon_r = 2.32$  dielectric. S parameter computations were run for various current expansion mode arrangements and up to 23 modes were used to check for convergence. Five or six piecewise modes on each side were found to be sufficient. The magnitude results are plotted in figure 6. These agree to some extent with quasistatic results for the thin dielectric limit. Phase calculations do not agree with quasistatic results. It is suspected that transverse directed current expansion modes will be necessary. Intuitively one might expect that the current must spread going from microstrip of one width to microstrip of another width. Thus it seems likely that a transverse current could be important.

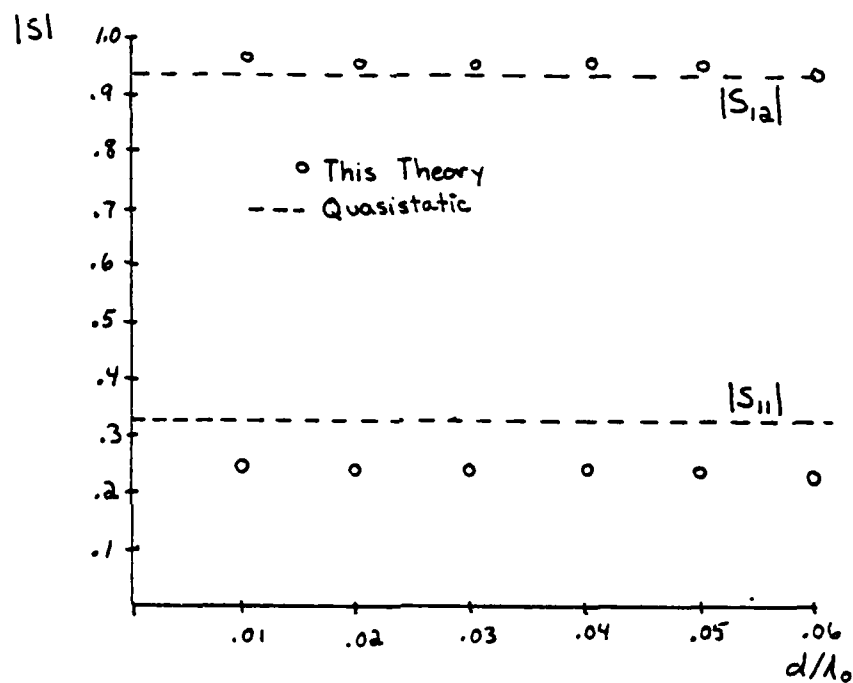


Figure 6. S parameter Magnitudes. Step Discontinuity  $100\Omega$ -to- $50\Omega$ .

The routine was checked by computing a through line (50 ohm to 50 ohm). This came out as expected (negligible  $S_{11}$  and  $S_{22}$ ;  $S_{21}$  and  $S_{12}$  equal 1). As another check, the open ended microstrip was calculated by selecting the elements of the impedance matrix which were formed only from + side currents and testing functions. These and + side  $V_j$  elements were used to form a mini matrix equation, the solution of which was compared to known good results for an open ended microstrip.

#### VI. Recommendations

In summary, results for the microstrip gap clearly show increasing loss to surface waves and radiation as substrate electrical thickness is increased. Step discontinuity calculations indicate increased loss with substrate thickness but phase calculations do not agree with quasistatic results. Transverse currents are probably necessary to accurately model this discontinuity.

It is recommended that the step discontinuity be reformulated to include transverse currents. Not only would this be likely to correct step phase modeling problems but it would be a very useful precursor to modeling the T junction and the right angle corner. The T junction especially is of interest to microwave circuit designers since the quasistatic model for the T junction is very suspect on thick substrates. A Research Initiation grant should include the investigation of these discontinuities.

The second recommendation involves investigation of alternate planar microwave and millimeter wave transmission line structures. The common ones include coplanar waveguide and slotline. The very useful part of the present formulation is that the programs written for microstrip discontinuities can be used to investigate discontinuities in these alternate structures. The only major change in any of the routines would be a change in the Green's function. Of course, convergence would have to be carefully checked but a great deal of time would be saved by not having to rewrite and debug a large amount of computer code.

#### ACKNOWLEDGEMENT

The author acknowledges the support of the Air Force Systems Command, the Air Force Office of Scientific Research and Rome Air Development Center, Electromagnetic Sciences Division, Hansom Air Force Base.

The author especially thanks Dr. John K. Schindler, Dr. Paul H. Carr and Andrew J. Slobodnik for their encouragement and help throughout this research. Finally, the author acknowledges many very helpful discussions with Professor David M. Pozar at the University of Massachusetts.

### References

- [1] "Capacitance Parameters of Discontinuities in Microstriplines", A. Gopinath and C. Gupta, MIT-25, Oct. 1977, pp. 819-822.
- [2] "Equivalent Circuit Parameters of Microstrip Step Change in Width and Cross Junctions", A. Gopinath, et.al., MIT-24, March 1976, pp. 124-144.
- [3] "Planar Waveguide Model for Calculating Microstrip Components", G. Komba and R. Mehran, Electron. Lett., Vol. 11, 1975, pp. 459-460.
- [4] "Full-Wave Analysis of an Open-Ended Microstripline", D. M. Pozar, submitted to IEEE Microwave Theory Tech.
- [5] D. M. Pozar, private communication.
- [6] "Input Impedance and Mutual Coupling of Rectangular Microstrip Antennas", D. M. Pozar, IEEE Trans. Antennas and Prop., VL. AP-29, pp. 99-105, Jan. 1981.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

ADVANCED PHYSICAL MODELING/WEDGE TEST DEVELOPMENT

Prepared by:	Vinod K. Jain, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Mechanical Engineering Department University of Dayton, Dayton, Ohio
Research Location:	Air Force Wright Aeronautical Laboratory Metals and Ceramics Division Materials Branch
USAF Research:	Dr. Harold L. Gegel
Date:	July 30, 1984
Contract No.:	F49620-82-C-0035

# ADVANCED PHYSICAL MODELING/WEDGE TEST DEVELOPMENT

by

Vinod K. Jain

## ABSTRACT

When an analytical model is developed to predict the behavior of a physical system, it is necessary to check the validity of the model by experimental procedure. An effort was directed to develop a laboratory test, the wedge test, to the level of a standard test to verify the analytical results of the ALPID computer program, which was designed to simulate the metal flow in deformation processing. The test employs a wedge-shaped specimen which is compressed in a die to study 2-dimensional and 3-dimensional deformation processes.

Tooling was designed for the 2-dimensional and 3-dimensional wedge tests, and the dies are currently being fabricated. A grid pattern, to be printed on the meridian plane of the specimen by a photochemical process, was developed. An analytical formulation has been developed to compute the various strain components using the dimensions of a deformed and undeformed grid.

#### ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with this excellent opportunity to work at the Air Force Wright Aeronautical Laboratory/Materials Laboratory, Wright-Patterson AFB, Dayton, Ohio. He would like to thank Dr. Harold L. Gegel for suggesting this area of research and for his guidance. Thanks are due to D. R. Barker, J. C. Malas, S. M. Doraivelu, L. E. Matson, Y. V. R. K. Prasad, and R. Srinivasan for their discussion and helpful suggestions during the course of this research.

## 1. INTRODUCTION

The mechanical behavior of materials under deformation processing is generally characterized by constitutive equations which relate the flow stress to strain, strain rate, and working temperature. Recently, a method for modeling the metal working processes has been developed, which uses a rigid-viscoplastic finite element method (FEM).<sup>1,2</sup> It predicts deformation behavior at selected points in each element (node point) by the application of a variational principle and gives an admissible solution using the initial boundary conditions. Whenever an analytical model is used to describe a system, a number of assumptions are always made. The latter affect the degree of accuracy of the final results. It is necessary that validity of the theoretical model be tested through physical modeling techniques. This provides feedback that aids in the refinement of the model and demonstrates the accuracy of the predictions.

For material deforming processes, two methods can be used to design the physical models. The first is to simulate, as closely as possible, a production metal-forming process, viz., extrusion or forging, on a laboratory scale and monitor the important parameters such as force, temperature, and ram velocity. The data from such a test can be analyzed to determine how closely its characteristics compare with those predicted by an analytical model. The second approach (to physical modeling) is to devise a laboratory test that is faster, easier, and less expensive than a subscale production test, yet tests the accuracy of the analytical model. These requirements are met by a test commonly used by the metal forming industry and is known as the wedge test. The test employs a wedge-shaped specimen as shown in Figure 1. The plane-strain deformation state can be obtained and studied by compressing the specimen in a channel-type die (Figure 2). Here the specimen is split and grid lines placed on the meridian plane to determine the strain components at any point resulting from the deformation. The 3-dimensional test is performed by compressing the specimen between a flat die and punch.

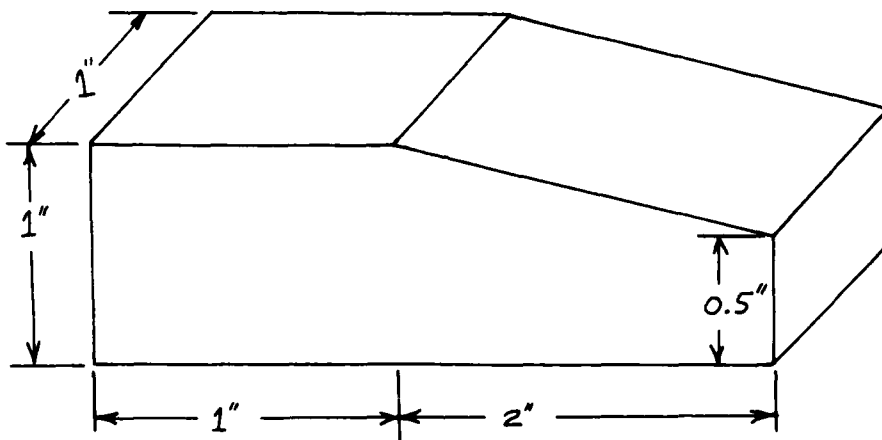


Figure 1. Schematic of a wedge specimen.

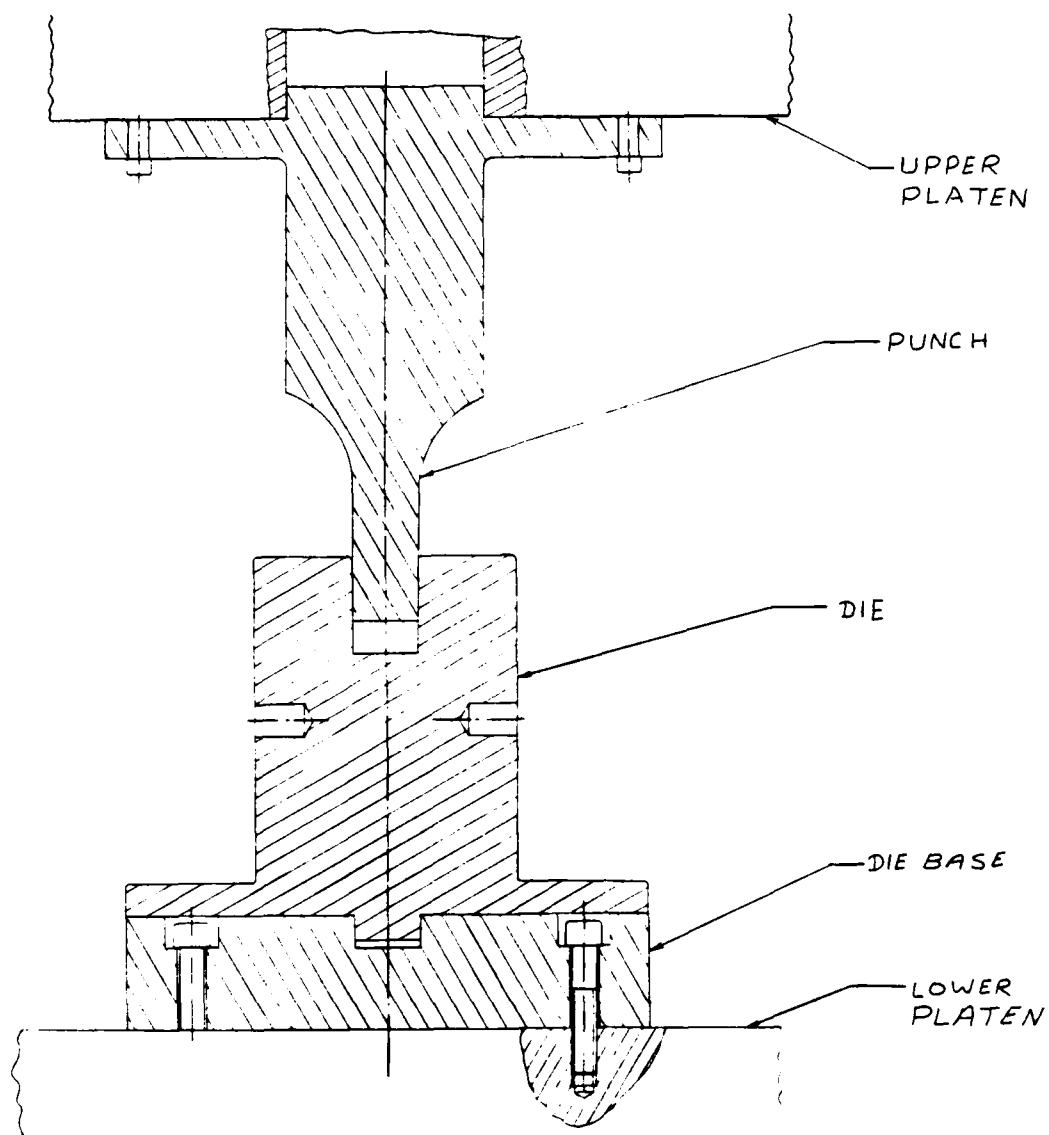


Figure 2. Layout of a punch and die for plane strain wedge test.

The Analytical Materials and Process Modeling Group at AFWAL/MLLM extensively uses an analytical model, Analysis of Large Plastic Incremental Deformation (ALPID), to design the extrusion dies and analyze the metal deformation, e.g., strain, strain rate, etc., during processing. It was therefore required to develop a physical test to verify the results of this computer program.

## 2. OBJECTIVES OF THE RESEARCH EFFORT

The chief objective of this work was to develop a laboratory test (the wedge test) to verify the results of the analytical modeling process used to evaluate the dynamic behavior of materials.

The following goals were initially set to accomplish the above objective:

- (a) To design and fabricate the tooling for both 2-d and 3-d deformation testing of wedge specimens.
- (b) To design a computerized data acquisition system capable of monitoring force, ram speed, and furnace temperature.
- (c) To perform the finite element analysis for the wedge-shaped specimen using the ALPID computer program.
- (d) To develop constitutive equations and perform power dissipation analysis for an Al-Fe-Ce system.
- (e) To compare the analytical results obtained using ALPID with those computed by physical modeling techniques.

During the course of research, steps (c) and (d) were dropped and the following additional goals were included:

- (A) To develop a grid pattern pertinent to the case of the wedge test and a technique to print it on aluminum surfaces.
- (B) To develop a mathematical analysis, using the fundamental definition of strain, to calculate the various strain components from the dimensions and geometry of the deformed and undeformed grids.

(C) To compute the strain distribution for 2-d and 3-d deformation testing of the wedge specimens using the mathematical analysis developed in (B) above.

### 3. APPROACH

The wedge test is unique since it results in different strains and microstructure at different points of the specimen. Thus, a variety of microstructures can be obtained in one single test.

#### 3.1 Design of Tooling

Tooling was designed to perform the wedge test to obtain 2-d and 3-d deformation of grid elements. The tests were to be performed on a 200 kip MTS tensile testing machine. The layout of the die and punch to be used with this machine for plane strain deformation is shown in Figure 2. The detail and layout drawings for both types of toolings (2-d and 3-d tests) were sent to the 4950th Test Wing for fabrication. Unfortunately, the dies were not ready, hence no experimental work could be carried out.

#### 3.2 Development of the Grid Mesh

A grid pattern is used to study the material flow in deforming solids. Some of the commonly accepted practices include: to scribe, emboss, etch, or photograph the grid pattern on the surface of the deforming solid.<sup>3</sup> Here, the purpose of the grid pattern is to study the metal flow in deformation process and to use the dimensions of the deformed and undeformed mesh to calculate the strain components. A chequerboard type of pattern, to be printed on the meridian plane of the wedge specimen (Figure 3), was developed for the plane strain deformation. It may be noted that the pattern consists of squares and parallelograms, whose shape (parallelograms) vary from point to point with the geometry of the specimen. The idea of using squares all over the surface was abandoned because it resulted in triangular grids at a number of locations and was not consistent with the grid mesh used by the ALPID. The original pattern was drawn at ten times



the full size and later reduced photographically to minimize the dimensional errors. The grid was printed on an aluminum sheet supplied by the Metal Photo Co., Cleveland, Ohio. Its thermal stability was checked, and the printed grid can hold up to 500°C without degrading or considerably fading away. An effort is now being made to have it printed on 0.5 in. thick sheet stock, which is not available from the above company.

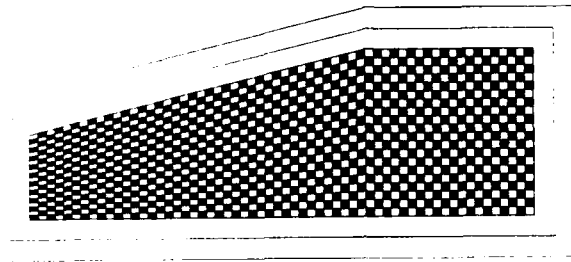


Figure 3. Chequerboard type of grid pattern for 2-d test.

#### 4. CALCULATION OF STRAIN FROM THE DIMENSIONS OF THE GRID MESH

In metal processing, it is required to evaluate the magnitude of strains in the deformed body. The formulas defining the strains are put in terms of the position of an element before and after the deformation with reference to some coordinate system. Thus, the strains are computed by measuring the deformation of some geometrical pattern involving a square grid mesh or an array of circles. The choice of pattern depends upon various factors, such as the type and accuracy of the information sought, material, and extent of the deformation of the specimen. In most plastic deformation processes the resulting strain is not homogeneous. The deformation however is assumed to be homogeneous. The rationale behind this assumption

is that the deformed material can be divided into a number of small regions where the deformation is homogeneous.

The most common method which is used to compute strains is called the visioplasticity.<sup>4-7</sup> It involves establishing a velocity vector field to obtain the strain. A grid pattern is placed on the meridian plane of a specimen, or at a plane right angle to metal movement, and the rate of the grid distortion is observed. The plane on which the grid is placed is photographed after each incremental deformation step, and the particle movement as represented by the intersection of the grid lines is observed. An adequate flow pattern is obtained in a few steps. Let  $V$  be the velocity of a grid point and its components in  $x$  and  $y$  directions are  $u$  and  $v$ , respectively. The values of  $u$  and  $v$  can be obtained for any desired point. The graphs,  $x$  vs  $v$ ,  $x$  vs  $u$ ,  $y$  vs  $u$ , and  $y$  vs  $v$  are then generated. Evaluation of slopes at each desired coordinate point from these curves permits the computation of strain rates as follows

$$\dot{\epsilon}_x = \frac{\partial u}{\partial x} \quad \dot{\epsilon}_y = \frac{\partial v}{\partial y} \quad \dot{\gamma}_{xy} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \quad (1)$$

The method is tedious and time consuming, and its accuracy depends on the computation of the velocity field. The technique is exact only in an engineering sense. It cannot be used for the wedge test because the deformation is unsymmetrical and a reference coordinate system cannot be established. The other method which involves the measurement of the dimensions of deformed and undeformed grids was found more appropriate for this case since it is direct and easier to use.

Sowerby et al.<sup>8,9</sup> recently proposed an analysis to compute the principal stretches in a deformed grid. The formulas depend on the initial shape of the grid, i.e., orthogonal or nonorthogonal. The technique involves measuring the deformed and undeformed lengths and the included angle of a quadrilateral grid mesh. For an orthogonal grid (Figure 4), the principal stretches  $\lambda_{11}$  and  $\lambda_{22}$  are given as follows:

$$\lambda_{11}^2 = \frac{\lambda_a^2 + \lambda_b^2}{2} + \frac{\lambda_a^2 - \lambda_b^2}{2} \frac{1}{\cos 2\theta} \quad (2)$$

$$\lambda_{22}^2 = \frac{\lambda_a^2 + \lambda_b^2}{2} - \frac{\lambda_a^2 - \lambda_b^2}{2} \frac{1}{\cos 2\theta} \quad (3)$$

$$\tan 2\theta = - \frac{\lambda_a \lambda_b \cos \phi}{\lambda_a^2 - \lambda_b^2}$$

where  $\lambda_a = \frac{a'}{a}$

$\lambda_b = \frac{b'}{b}$

$\phi$  = angle between the deformed pair of grid lines

$\theta$  = angle between the principal direction 1 and x axis

$a$  = length of the undeformed line element OA

$a'$  = length of the deformed line element OA'

$b$  = length of the undeformed line element OC

$b'$  = length of the deformed line element OC'

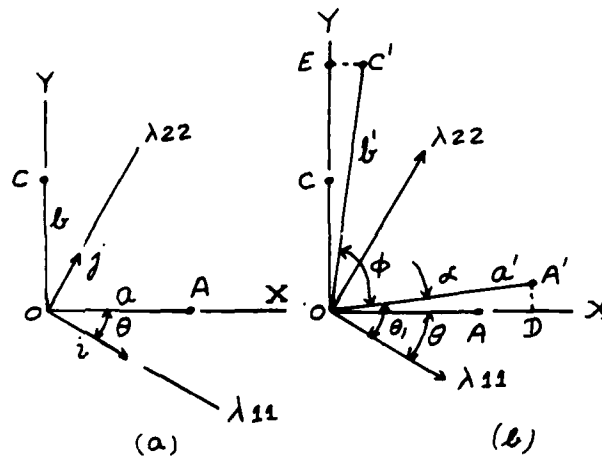


Figure 4. (a) Undeformed orthogonal line elements;  
(b) deformed state following homogeneous strain.

For a nonorthogonal grid (Figure 5), the principal stretches are given

by

$$\lambda_{11}^2, \lambda_{22}^2 = \frac{\frac{1}{2} \{ (\lambda_a^2 + \lambda_b^2) \cos 2\theta \pm (\lambda_a^2 - \lambda_b^2) \} - \lambda_a^2 \cos \xi \cos (\xi + 2\theta)}{\sin \xi \sin (\xi + 2\theta)} \quad (4)$$

where  $\xi$  = angle between the sides of an undeformed grid.

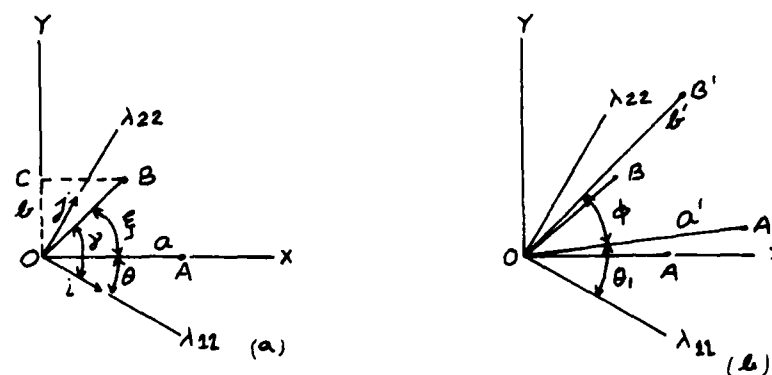


Figure 5. (a) Undeformed nonorthogonal line elements;  
(b) deformed state following homogeneous strain.

$$\cos \phi = \frac{1}{\lambda_a \lambda_b} (\lambda_{11}^2 \cos \theta \cos \gamma + \lambda_{22}^2 \sin \theta \sin \gamma)$$

$$\cot(\xi + 2\theta) = \frac{2 \lambda_a \lambda_b \cos \phi - (\lambda_a + \lambda_b) \cos \xi}{\sin \xi (\lambda_a^2 - \lambda_b^2)}$$

The parameters  $\lambda_a$ ,  $\lambda_b$ ,  $\theta$ , and  $\phi$  have the same meaning as before.

The above method is mostly recommended for orthogonal grids<sup>9</sup> and suffers from the drawback of calculating the angles  $\phi$  and  $\xi$  for every grid point (node).

In view of the above difficulties, a new scheme was devised to calculate the strains based on the coordinate displacements of an element. Since the actual strains are large, no limitations were applied to their magnitude and general equations of finite homogeneous strains in terms of coordinate displacement developed.

#### 4.1 Normal Strains

Consider an element, as shown in Figure 6, of a body whose sides are parallel to cartesian coordinates  $x$ ,  $y$ ,  $z$  in the undeformed state. The sides of the element are of lengths  $\delta x$ ,  $\delta y$ , and  $\delta z$ . Thus, the coordinates of A are  $x$ ,  $y$ ,  $z$ , and that of B ( $x + \delta x$ ), ( $y + \delta y$ ), and ( $z + \delta z$ ). In the deformed condition, the whole element is displaced owing to the accumulated strains. The point A moves to A' and B to B'. If the displacement components of A' are  $u$ ,  $v$ , and  $w$ , then its coordinates will be  $(x + u)$ ,  $(y + v)$ , and  $(z + w)$ . Each side of the element would be strained, e.g., AC deforms to A'C'. The latter may be due to a combination of normal and shear strains. It is assumed that the element AC is very small and that the changes in  $\delta x(AC)$  are proportional to its length. Therefore, the length of the  $x$  component of A'C' is  $\delta x + \frac{\partial u}{\partial x} \delta x$ ,  $y$  component is  $\frac{\partial v}{\partial x} \delta x$ , and the component parallel to the  $z$  direction is  $\frac{\partial w}{\partial x} \delta x$  (Figure 7).

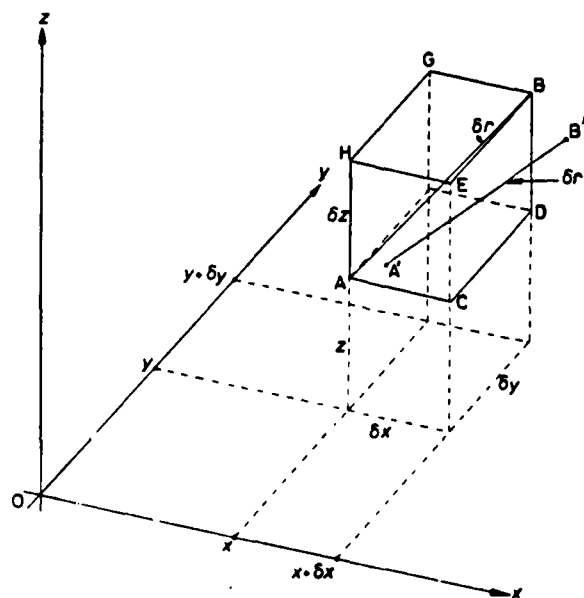


Figure 6. Schematic of a 3-d element (Ref. 10).

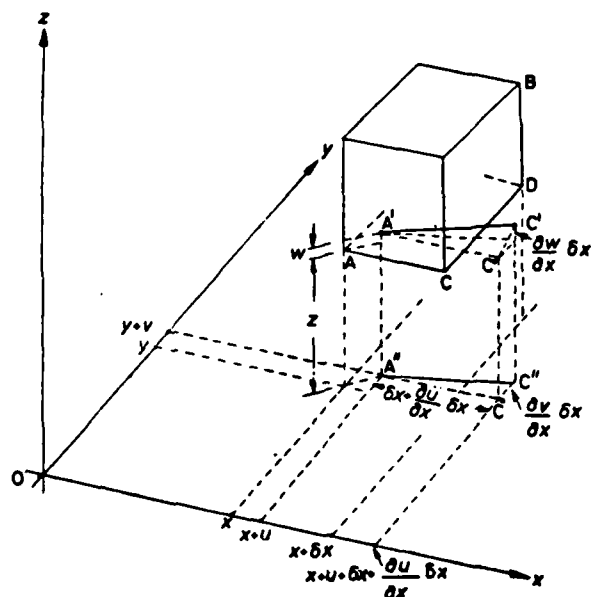


Figure 7. Deformation of a line element due to normal and shear strains (Ref. 10).

Then

$$(A'C')^2 = (\delta x + \frac{\partial u}{\partial x} \delta x)^2 + (\frac{\partial v}{\partial x} \delta x)^2 + (\frac{\partial w}{\partial x} \delta x)^2$$

$$(A'C')^2 = \delta x^2 \{ 1 + 2 \frac{\partial u}{\partial x} + (\frac{\partial u}{\partial x})^2 + (\frac{\partial v}{\partial x})^2 + (\frac{\partial w}{\partial x})^2 \} \quad (5)$$

If

$$e_{xx} = \frac{\partial u}{\partial x} + \frac{1}{2} \{ (\frac{\partial u}{\partial x})^2 + (\frac{\partial v}{\partial x})^2 + (\frac{\partial w}{\partial x})^2 \}$$

Then, according to the general definition, strain,  $e_x$ , in AC is

$$e_x = \frac{A'C'}{AC} - 1 = \sqrt{1 + 2 e_{xx}} - 1 \quad (6)$$

Similarly, expressions for  $e_{yy}$  and  $e_{zz}$  can be written as

$$e_{yy} = \frac{\partial v}{\partial y} + \frac{1}{2} \{ (\frac{\partial u}{\partial y})^2 + (\frac{\partial v}{\partial y})^2 + (\frac{\partial w}{\partial y})^2 \} \quad (7)$$

$$e_{zz} = \frac{\partial w}{\partial z} + \frac{1}{2} \{ (\frac{\partial w}{\partial z})^2 + (\frac{\partial v}{\partial z})^2 + (\frac{\partial u}{\partial z})^2 \} \quad (8)$$

where the actual strains  $e_y$  and  $e_z$  are

$$e_y = \sqrt{2 e_{yy} + 1} - 1 \quad (9)$$

and

$$e_z = \sqrt{1 + 2 e_{zz}} - 1 \quad (10)$$

#### 4.2 Shear Strains

The shear strains of the element AB (Figure 6) are the changes in angle or distortions for the point A, i.e., the changes in three right angles at the corner A. The shear strain in the  $xoy$  plane, as shown in Figure 8, is given by the change in angle HAC,<sup>10</sup>

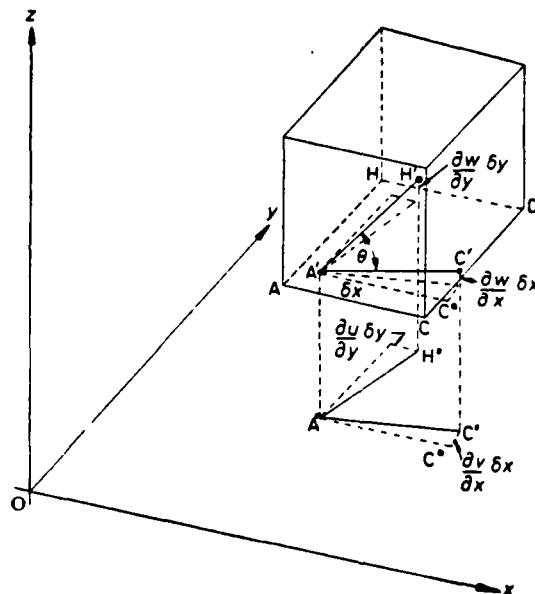


Figure 8. Deformation of an element due to shear strains.

that is  $(\frac{\pi}{2} - \theta)$ , where the angle  $H'A'C'$  in the distorted plane is  $\theta$ . The shear strain,  $e_{xy}$ , in the plane  $xoy$  can be shown to be

$$e_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial w}{\partial x} \frac{\partial w}{\partial y} \quad (11)$$

Similarly, expressions for the shear strains in  $yoz$  and  $zox$  planes can be written as

$$e_{yz} = \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} + \frac{\partial u}{\partial y} \frac{\partial u}{\partial z} + \frac{\partial v}{\partial y} \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \frac{\partial w}{\partial z} \quad (12)$$

and

$$e_{zx} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial z} \frac{\partial v}{\partial x} + \frac{\partial w}{\partial z} \frac{\partial w}{\partial x} \quad (13)$$

#### 4.3 Strain in Any Direction

Using the above equations, a general expression of strain for an element in any direction, e.g., AB in Figure 6, will be derived. Let  $\delta r$  be the undeformed length of AB which after deformation moves to A'B' having a length  $\delta r'$ . If  $l_r$ ,  $m_r$ , and  $n_r$  are the direction cosines of the element  $\delta r$ , and  $l'_r$ ,  $m'_r$ , and  $n'_r$  of the strained element  $\delta r'$ , then

$$l_r = \frac{\delta x}{\delta r}, \quad m_r = \frac{\delta y}{\delta r}, \quad n_r = \frac{\delta z}{\delta r},$$

$$\text{and } l'_r = \frac{\delta x + \delta u}{\delta r'}, \quad m'_r = \frac{\delta y + \delta v}{\delta r'}, \quad n'_r = \frac{\delta z + \delta w}{\delta r'}$$

$$\delta u = \frac{\partial u}{\partial x} \delta x + \frac{\partial u}{\partial y} \delta y + \frac{\partial u}{\partial z} \delta z$$

$$\delta v = \frac{\partial v}{\partial x} \delta x + \frac{\partial v}{\partial y} \delta y + \frac{\partial v}{\partial z} \delta z$$

$$\delta w = \frac{\partial w}{\partial x} \delta x + \frac{\partial w}{\partial y} \delta y + \frac{\partial w}{\partial z} \delta z$$

Then

$$l'_r = \frac{\delta x}{\delta r} \frac{\delta r}{\delta r'} + \frac{\delta r}{\delta r'} \frac{\delta u}{\delta r}$$

$$\text{or } l'_r = \frac{\delta r}{\delta r'} \left\{ l_r \left( 1 + \frac{\partial u}{\partial x} \right) + m_r \frac{\partial u}{\partial y} + n_r \frac{\partial u}{\partial z} \right\} \quad (14)$$

Since  $l'^2_r + m'^2_r + n'^2_r = 1$ , squaring and adding equation (14) and the other corresponding expressions result in

$$\begin{aligned} \left( \frac{\delta r'}{\delta r} \right)^2 &= (1 + 2 e_{xx}) l_r^2 + (1 + 2 e_{yy}) m_r^2 + (1 + 2 e_{zz}) n_r^2 + 2 e_{xy} l_r m_r \\ &\quad + 2 e_{yz} m_r n_r + 2 e_{zx} n_r l_r \end{aligned} \quad (15)$$



For the case of plane strain deformation as in the 2-d wedge test,  
 $e_{yz} = e_{zx} = e_{zz} = n_r = 0$ , and equation (15) reduces to

$$\left(\frac{\delta r'}{\delta r}\right)^2 = (1 + 2e_{xx}) l_r^2 + (1 + 2e_{yy}) m_r^2 + 2e_{xy} l_r m_r \quad (16)$$

#### 5. APPLICATION OF THE GENERAL STRAIN EQUATION TO THE CASE OF PLANE STRAIN DEFORMATION

For the case of plane strain deformation, the strain in  $z$  direction is zero. Consider an orthogonal element (Figure 9a) OABC which after deformation takes the shape OA'B'C' (as shown in Figure 9b).

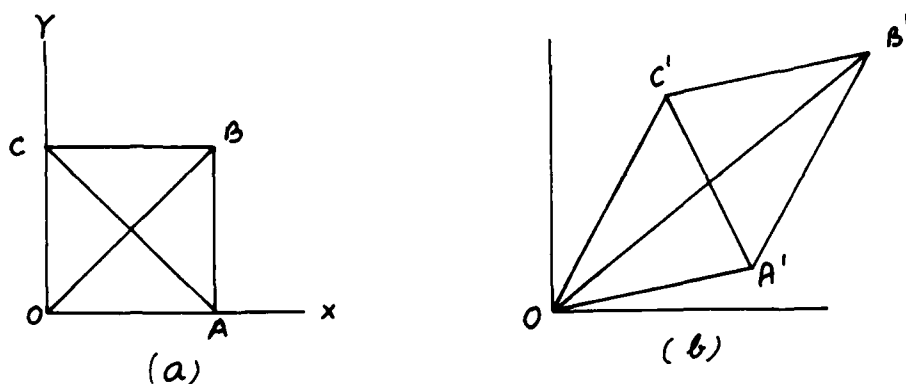


Figure 9. (a) Undeformed grid; and  
 (b) deformed grid following homogeneous strain.

Let the direction cosines of OB and AC be  $l_1, m_1$ , and  $l_2, m_2$  respectively. The lengths of OB and AC can be conveniently computed with the digitized data of the deformed and undeformed grids. Equation (16) for the elements OB and AC reduces to

$$\left(\frac{OB'}{OB}\right)^2 = (1 + 2e_{xx}) l_1^2 + (1 + 2e_{yy}) m_1^2 + 2e_{xy} l_1 m_1 \quad (17)$$

or

$$\left(\frac{OB'}{OB}\right)^2 = 1 + 2l_1^2 e_{xx} + 2m_1^2 e_{yy} + 2e_{xy} l_1 m_1$$

Similarly

$$\left(\frac{A'C'}{AC}\right)^2 = 1 + 2 \ell_2^2 e_{xx} + 2 m_2^2 e_{yy} - 2 e_{xy} \ell_2 m_2 \quad (18)$$

For plane-strain plastic deformation

$$\epsilon_x = -\epsilon_y$$

$$\text{or} \quad \ln(1+e_x) = -\ln(1+e_y) = \ln \frac{1}{1+e_y}$$

$$\text{or} \quad 1+e_x = \frac{1}{1+e_y}$$

$$\text{But} \quad e_x = \sqrt{1+2e_{xx}} - 1$$

$$\text{Then} \quad 1+2e_{xx} = \frac{1}{1+2e_{yy}}$$

$$\text{or} \quad e_{yy} = -\frac{e_{xx}}{1+2e_{xx}} \quad (19)$$

Substitution of equation (19) in (17) gives

$$\begin{aligned} \left(\frac{OB'}{OB}\right)^2 &= 1 + 2 \ell_1^2 e_{xx} - 2 m_1^2 \frac{e_{xx}}{1+2e_{xx}} + 2 e_{xy} \ell_1 m_1 \\ \text{or} \quad e_{xx}^2 (4 \ell_1^2) &+ 2 e_{xx} \left[ \left(1 - \left(\frac{OB'}{OB}\right)^2\right) + \ell_1^2 - m_1^2 \right] + \\ &2 e_{xy} \ell_1 m_1 (1+2e_{xx}) + \left[1 - \left(\frac{OB'}{OB}\right)^2\right] = 0 \end{aligned} \quad (20)$$

$$\begin{aligned} \text{Let} \quad A_1 &= 4 \ell_1^2 \\ B_1 &= 2 \left[ \left(1 - \left(\frac{OB'}{OB}\right)^2\right) + \ell_1^2 - m_1^2 \right] \\ C_1 &= 2 \ell_1 m_1 \\ D_1 &= 1 - \left(\frac{OB'}{OB}\right)^2 \end{aligned} \quad (21)$$

Substitution of equation (21) in (20) gives

$$A_1 e_{xx}^2 + B_1 e_{xx} + C_1 (1+2e_{xx}) e_{xy} + D_1 = 0 \quad (22)$$

Similarly, equation (18) can be written as

$$A_2 e_{xx}^2 + B_2 e_{xx} - C_2 (1 + 2 e_{xx}) e_{xy} + D_2 = 0 \quad (23)$$

To eliminate  $e_{xy}$  from the above equations, multiply equation (22) by  $C_2$  and equation (23) by  $C_1$  and add the resulting expressions

$$(A_1 C_2 + A_2 C_1) e_{xx}^2 + (B_1 C_2 + B_2 C_1) e_{xx} + (D_1 C_2 + D_2 C_1) = 0 \quad (24)$$

Again, let

$$A' = A_1 C_2 + A_2 C_1$$

$$B' = B_1 C_2 + B_2 C_1$$

$$D' = D_1 C_2 + D_2 C_1$$

Equation (24) can now be simplified as

$$A' e_{xx}^2 + B' e_{xx} + D' = 0 \quad (25)$$

The quadratic equation (25) will result in two values of  $e_{xx}$ , one of which will be negative and the other positive. The negative value would be inadmissible since the strain in the x direction cannot be compressive. The value of  $e_{xy}$  can now be calculated from equation (22) or equation (23). Actual values of strains  $e_x$  and  $e_y$  are determined from

$$e_x = \sqrt{1 + 2 e_{xx}} - 1 \quad (26)$$

and

$$e_y = - \frac{e_x}{1 + e_x} \quad (27)$$

The other relationships used to calculate the various strain components are

$$\text{Effective Strain, } \bar{e} = \frac{2}{3} \sqrt{3 e_x^2 + \frac{3}{4} e_{xy}^2} \quad (28)$$

$$\text{Effective True Strain, } \bar{\epsilon} = \ln (1 + \bar{e}) \quad (29)$$

Principal  
Strains,

$$e_1, e_2 = \frac{e_x + e_y}{2} \pm \sqrt{\left(\frac{e_x - e_y}{2}\right)^2 + \frac{1}{4}(e_{xy})^2} \quad (30)$$

The true principal strains  $\epsilon_1$  and  $\epsilon_2$  are given by

$$\epsilon_{1,2} = \ln(1 + e_{1,2}) \quad (31)$$

Thus, the above analysis can be used to determine the various strain components in a deformed body without making any assumption about their magnitudes.

#### 6. COMPUTATION OF STRAIN COMPONENTS

The equations derived earlier were used to determine the various strain components using a deformed and undeformed grid. Unfortunately, actual tests could not be performed since the dies were not fabricated as on this day of writing the report. Hence, an academic guess was made, and the deformed grid pattern was drawn as shown (at a reduced scale) in Figure 10. Figure 10a shows the line sketch of the grid pattern shown in Figure 3. An element OABC was selected for the calculation of strain components.

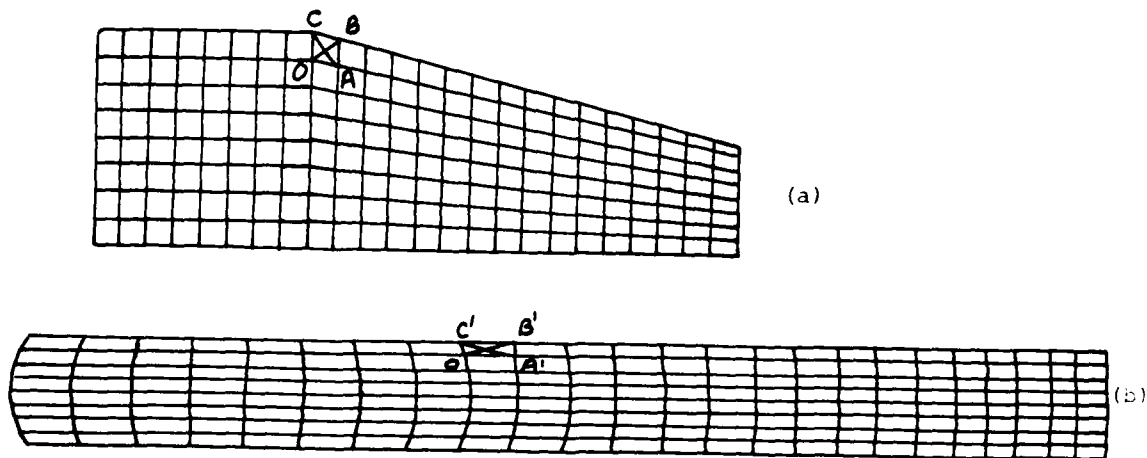


Figure 10. Schematic of the grid pattern for the wedge test:  
(a) undeformed; (b) deformed after the test.

The following dimensions were obtained by the measurement of the line elements of deformed and undeformed grid OABC;

$$\begin{aligned} AC &= 0.451 & \ell_1 &= 0.80 \\ A'C' &= 0.531 & m_1 &= 0.80 \\ OB &= 0.3125 & \ell_2 &= 0.554 \\ OB' &= 0.439 & m_2 &= 0.831 \end{aligned}$$

Substitution of the values of OB, OB', AC, A'C',  $\ell_1$ ,  $\ell_2$ ,  $m_1$ , and  $m_2$  in equations (25) to (31) resulted in the following strain values:

$$e_{xx} = 1.134, e_{xy} = 0.067, e_x = .8077, e_y = -.4468$$

$$\bar{e} = 0.9381, \bar{\epsilon} = 0.66, e_1 = 0.808, e_2 = -0.448$$

$$\epsilon_1 = 0.592, \epsilon_2 = -0.594, \epsilon_x = 0.592, \epsilon_y = -0.592$$

It may be noted that the above results satisfy the condition of constant volume for plastic deformation, that is

$$\epsilon_1 + \epsilon_2 = 0$$

$$\epsilon_x + \epsilon_y = 0$$

## 7. RECOMMENDATIONS

The objective of this research work was to develop a laboratory test (the wedge test) to the level of a standard test to compare the results of the ALPID computer program. Unfortunately, the work could not be completed for two reasons. The first is that the tooling required for the experiments could not be fabricated in time, hence the experiments could not be performed. The second is that the problem is too large to be completed in ten weeks time. In order to accomplish the stated objectives of this research, additional work as outlined below should be conducted:

1. Print grid pattern on 2-dimensional wedge specimens to study the plain strain deformation process.

2. Develop a grid system for 3-dimensional testing of the wedge specimen.
3. Digitize the undeformed grid pattern.
4. Test the wedge specimens at several ram speeds and temperatures.
5. Digitize the deformed grid for each test specimen.
6. Develop a computer program to calculate the various strain components using the digitized data, and draw contour maps for them.
7. Compare the wedge test results with those of ALPID results.
8. Submit a final report on the above research work.

#### 8. REFERENCES

1. C. H. Lee and S. Kobayashi, "New Solutions to Rigid Plastic Deformation Problems Using a Matrix Method," Trans. ASME, J. Engr. Ind., 95, p. 865 (1973).
2. S. I. Oh, "Finite Element Analysis of Metal Forming Problems with Arbitrarily Shaped Dies," Int. J. Mech. Sci., 24, p. 479 (1982).
3. L. E. Farmer and R. F. Fowle, "An Experimental Procedure for Studying the Flow in Plane Strain Extrusion," Int. J. Mech. Sci., 21, p. 599 (1979).
4. E. G. Thomsen, C. T. Yang, and S. Kobayashi, Plastic Deformation in Metal Processing, New York, The Macmillan Co. (1965).
5. A. H. Shabaik and E. G. Thomsen, "A Theoretical Method for the Analysis of Metal-Working Problems," J. of Eng. for Ind., p. 343 (May 1968).
6. A. H. Shabaik and E. G. Thomsen, "Computer Application to the Visioplasticity Method," J. of Eng. for Ind., p. 343 (May 1967).
7. L. E. Farmer and P. L. B. Oxley, "A Computer-Aided Method for Calculating the Distributions of Strain Rate and Strain from an Experimental Flow Field," J. of Strain Analysis, 11, p. 26 (1976).
8. R. Sowerby, E. Chu, and J. L. Duncan, "Determination of Large Strains in Metal Forming," J. of Strain Analysis, 17, p. 95 (1982).
9. R. Sowerby and P. C. Chakravarti, "The Determination of the Equivalent Strain in Finite, Homogeneous Deformation Process," J. of Strain Analysis, 18, p. 121 (1983).
10. H. Ford, Advanced Mechanics of Materials, New York, Halsted Press (1977).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DEVELOPMENT OF A VEHICLE FLEET ANALYSIS SYSTEM

Prepared by:	Bruce N. Janson
Academic Rank:	Assistant Professor
Department and University:	Department of Civil Engineering Carnegie-Mellon University
Address:	Pittsburgh, PA 15213
Research Location:	Air Force Logistics Management Center, Gunter AFS
USAF Research:	Capt. James R. Van Scotter, USAF
Date:	August 1, 1984
Contract No:	F49620-82-C-0035

## DEVELOPMENT OF A VEHICLE FLEET ANALYSIS SYSTEM

by

Bruce N. Janson

### ABSTRACT

This report documents the functional specification, design and development of a prototype vehicle fleet analysis system. To facilitate modular development, the system was divided into three functional areas: mainframe-to-micro communication, database management, and fleet analysis capabilities. The eventual system will be comprised of a set of micro-computer based modules that will enable more effective utilization of transportation resources by Air Force vehicle fleet managers and staff. Key requirements of the system configuration are that:

- a. A smooth communication link between the micro-computer and the mainframe Vehicle Integrated Management System must be established.
- b. Database structures must be compatible with each functional area of the system.
- c. The system must be highly interactive and instructive to the user.
- d. The system must be entirely menu driven and capable of performing most individual requests within seconds.

Since efficient mainframe-to-micro communication had not yet been established, one major task of the project was to compile a sample database for testing prototype modules as they are developed. Monthly data from Randolph Air Force Base was used for this purpose. A final major task of the 10-week research period was to develop an initial data display and graphical analysis demonstration module. This program was used in later conferences and roundtable discussions to focus attention on potential capabilities, user needs and technical considerations for a system of this type. Recommendations for how to best continue the development process and important research areas to investigate are made.



AD-A154 337

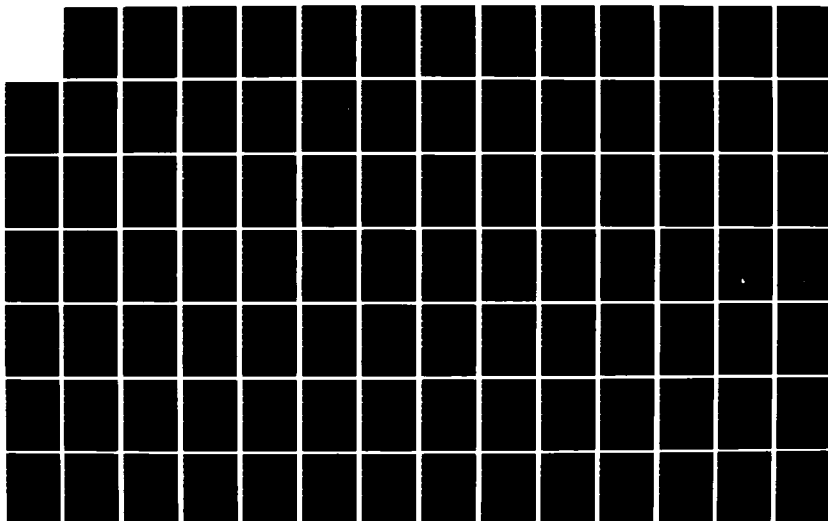
UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 6/13  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

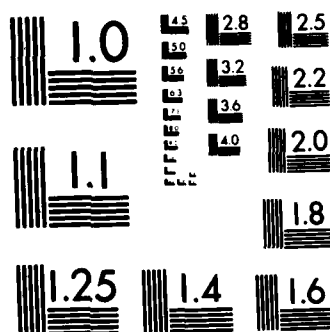
UNCLASSIFIED

M D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

## 1. INTRODUCTION

The U.S. Air Force owns and operates over 100,000 ground vehicles at major and minor installations across the world. These vehicles are essential to support daily base-level operations and, perhaps more critically, for special mission support. This fleet is composed of roughly 330 different types of vehicles ranging from sedans and pick-up trucks to buses, fire trucks and heavy construction equipment. Due to the continuously changing age, composition and condition of this fleet, it is imperative that better vehicle fleet management technology be developed to assist vehicle fleet managers in providing ground support to the Air Force in a responsive, yet cost effective, manner.

The Air Force has set out requirements for quarterly and annual analyses of vehicle fleet operations and maintenance in Volumes I and II of Air Force Manual 77-310. [1] However, there is insufficient time to perform these analyses by hand given the other competing activities and administrative duties of an Air Force fleet manager and staff. More importantly, the time and effort required to compile and analyze vehicle use data by hand obstructs the fleet manager from evaluating the fleet's current utilization and performance in light of the demands for services being placed upon it.

Through this research, we set out to develop a prototype fleet analysis system comprised of a set of micro-computer based programs or modules that would enable more effective management of resources by base-level fleet managers and staff. This system, as outlined in the following sections, is required to be highly interactive and instructive to the user, as well as being adaptable to a wide variety of user needs. A system development effort can often gain or lose ground depending upon the needs of the user, either stated by the user or assumed by developers of the system. [2] This report summarizes the work we accomplished towards developing this system during a 10-week research engagement with the Air Force Logistics Management Center (AFLMC) at Gunter Air Force Station in Alabama.

## 2. OBJECTIVES

The mission of the AFLMC is to study and seek solutions to base-level logistics problems. As such, the Transportation Directorate (LGT) at AFLMC

has conducted several projects to improve upon currently used vehicle fleet management techniques. A recent LGT project addressed the need to merge and consolidate the vehicle fleet management data found in two separate information systems: the Registered Equipment Management System (REMS) and the Vehicle Integrated Management System (VIMS). As a result of this project, the Air Staff approved the development of a prototype base-level vehicle fleet analysis system that would enhance the use of both REMS and VIMS data for fleet management and analysis purposes. The Transportation Directorate at AFLMC was directed to conduct this project as outlined by the Project Plan. [3]

Taking into account the 10-week period that I would be working at the AFLMC, the objectives of this study were laid out as follows:

1. develop a list of functional specifications detailing what information is expected from the fleet analysis system, and what types of analyses the system is expected to perform.
2. determine the best approach for system development, i.e. write new code from scratch or incorporate as much currently available Z-100 spreadsheet and database software as possible.
3. compile a sample database that can be used for testing prototype modules of the fleet analysis system as they are being developed, and transfer this data into the Z-100 microcomputer so that it can be readily accessed by the system.
4. subdivide the fleet analysis system into a series of modules that can be developed, demonstrated and then tested at selected bases so as to encourage feedback from Air Force personnel who will eventually be using the system.
5. develop the first set of data display modules that provide fleet managers with pertinent information in a clear, concise and responsive manner using menu driven requests for tabular and graphic displays of fleet utilization and performance statistics.

The above list of research tasks coincides closely with those proposed in the initial statement of work. Despite the brevity of the research period, we were able to accomplish most of the above tasks. In cases where certain difficulties, such as software availability or mainframe-micro communication, obstructed our progress on one task, we turned our attention to a different task so as not to delay the entire project. It was not possible in this short period of time to encourage the degree of user involvement that is known to be critically important to the success of a system of this type. [4] We were,

however, able to receive a great variety of inputs and critical reviews at the Vehicle Fleet Analysis Conference. Our accomplishments to date for each of the above taskings are described in the remainder of this report.

### 3. FUNCTIONAL SPECIFICATIONS

The first and most important task of any system development project is to outline, in as much detail as possible, just how the system is expected to perform. These specifications must include not only what types of analyses the system must be able to perform at the user's request, but also how the results of those analyses ought to be formatted and displayed, and also with what ease and speed a trained user should be able to accomplish a given task. The system must be focused towards "user needs" and not simply "analysis needs" in order to be a functionally useful tool. [5] Since system requirements often give rise to competing objectives, we developed the following list of questions pertaining to system capabilities in order to guide our determination of functional specifications.

- a. What types of analyses must a user of the system be capable of performing (e.g. inquiries, updating, selective record retrieval, basic statistics, trend analysis, forecasting, etc.)?
- b. What types of output display are required (e.g. standard report tables, self-formatted tables, graphics options, screen vs. hardcopy display)? [6]
- c. How large a database must the system be able to handle, and what impact will the size of the database have on the system's computational speed?
- d. How must the user interface be designed (e.g. will the system be entirely menu driven or will certain stages of analysis require the user to enter explicit programming type commands)?
- e. Finally, will the user be expected to enter and maintain any of the data elements, or will all data elements be passed directly to the vehicle fleet analysis system via an automated data transfer link with the mainframe computer?

This last item is perhaps the most difficult to control in the entire project because the smoothness with which the fleet analysis system on a Z-100 micro-computer interfaces with the mainframe computer is dependent on current mainframe-to-micro communication technology and the extent to which we are assisted by other Air Force personnel having direct experience in using that

technology. Before describing that aspect of the project, a list of functional specifications that we agreed upon for an initial prototype system are described below.

- a. A user must be able to find, examine and retrieve any record or records that meet user specified criteria. Types of criteria that a user can specify should be made as flexible as possible, yet easy to enter.
- b. A user must be able to aggregate information by vehicle type automatically based upon the file of individual vehicle records present in the system. Inquiries to information by vehicle type as required above for individual vehicle records must also be possible via the same style of user commands.
- c. A user of the system must be able to create a tabular display of whatever data elements are requested for whatever records of information meet the specified criteria.
- d. A user must be able to graph, with reasonable flexibility any set of applicable data elements that have been identified through capabilities (a,b,c) above.
- e. A user must be able to view all output generated through capabilities (a,b,c,d) at the Z-100 monitor, with the option of creating hardcopy output if a printer is available.
- f. While basic statistics (sums, averages, ratios) of data by vehicle type may be sufficient information from a prototype system, the system must be designed such that new modules to perform trend analysis, linear regression, and forecasting can be smoothly interfaced with existing modules.
- g. Responsiveness of the system to user requests must be rapid enough so as to command the user's attention. Complex inquiries to a large database should require no more than 3 minutes of execution time by the Z-100. The system should execute most commands within 5 seconds, and require only 1 or 2 seconds for data entry and monitor control commands.
- h. The system must be capable of handling databases of up to 2500 individual vehicle records having as many as 50 data elements per record, either raw or calculated. Analysis modules can be designed to handle fewer vehicles and data elements at one time as required by the active working memory of the micro-computer. However, the extraction of a secondary database for analysis from the main database must be a smooth, efficient and logical process.
- i. The system must be able to receive data passed to it from the On-Line VIMS/REMS system soon to be installed on Air Force Sperry mainframe computers worldwide. This data communication link is vital to the success of an analysis system, since the duplication of data entry effort must be avoided at all costs. The analysis system must use

historical and current data in the VIMS/REMS system to the best extent possible, since gathering other data would be a major task. [2]

- j. The system must be entirely menu driven with the parallel availability of on-screen help files at every step. A user's manual that elaborates upon the help files and cryptic menu command explanations should accompany the system for detailed reference and study.
- k. Finally, the system must be capable of providing annual and quarterly analyses of vehicle fleet data as required by Volume I of Air Force Manual 77-310. Volume I states that vehicle fleet analysis should include:

- (i) by vehicle type
  - a. trends of vehicle utilization
  - b. operations and maintenance costs
  - c. fuel consumption per mile or hour
  - d. vehicle out-of-commission rates
- (ii) identification of causative factors and recommendations for improvement
- (iii) comparisons of the above to base-level averages and Air Force standards

Although system specifications listed above allow a for great deal of flexibility as to what data is included in the database, what calculations are performed, and just how the results are presented, the intent of these specifications is just that -- to create a flexible system structure that is able to accommodate both command and base-level requests for modifications and enhancements. Towards the end of the 10-week research period, the AFLMC held a Vehicle Fleet Analysis Conference attended by Air Force transporters from across the U.S. This conference resulted in a detailed list of analyses useful to fleet managers and how these analyses ought to be presented. The sheer bulk of tables and graphs representing their requests precludes their inclusion in a report of this size. Moreover, the conference identified several areas of debate among the attendees as to how best perform and present certain analyses. We feel certain that the automation of vehicle fleet analyses itself will induce even further add-ons to the range of analysis capabilities requested once the user is able to assess the degree of flexibility made possible through automation.

#### 4. SYSTEM DEVELOPMENT

After defining the system specifications, the next task was to determine the best approach for system development. The key decision here was whether the system should be coded from scratch in a higher level programming language such as PASCAL, or to incorporate as much currently available Z-100 spreadsheet and database software as possible. Key factors that impact this decision are:

- a. characteristics of the user interface (visual, responsive, instructive, menu driven).
- b. ease and speed of programming (development horizon and expertise required).
- c. taking advantage of inquiry, frame entry, graphics and report generation utilities already established in existing software packages.
- d. computational speed on large data files.
- e. random access (RAM) versus permanent (hard or floppy disk) memory constraints and requirements of the chosen system approach.
- f. capability of the system to adopt or be adapted to new modules or new hardware/software technology as required or deemed advantageous.

While each of the above had an impact on our development approach, some were considered more vital than others. For example, while computational speed is important, any system that provides several magnitudes of improvement over current by-hand methods would be welcomed graciously. Likewise, while programming the system from scratch in PASCAL may require reinventing certain user interface utilities, this approach may be the best in terms of optimizing computation speed and data storage requirements. In the end, several factors pertaining to user interfacing and capability of the system to adopt new requirements without a great deal of re-programming led us to adopt the approach of using as much existing software for the Z-100 as possible.

In deciding to incorporate existing software for the Z-100, our main concern was whether currently available software could handle a database of size 2500 x 50, our stated requirement. To address our concerns, I conducted a brief literature review of available software to compare their hardware requirements and performance indicators. An important conclusion reached from



this investigation is that we must be careful to select software that is best suited for each functional area of the fleet analysis system. Fortunately, the system decomposes into just three primary functional areas:

1. monthly retrieval of data from the mainframe computer; extensive computations can be performed in this data retrieval process, as well as formatting the data to be compatible input to the next stage of software.
2. automated and visual inquiry to specific records; selective record updating and retrieval; standard report generation and formatting.
3. special manipulation of selective data elements; optional analyses such as graphic displays, trend analysis, analysis of variance, regression, etc.

For reasons explained next, the above sub-division of system functions corresponds well to the appropriate uses of current micro and mainframe computer hardware/software technology. Functional Area 1 is best accomplished through the use of AFOLDS retrieval programs written in COBOL with additional calculations and database preparation routines built in. Functional Area 2 is best accomplished through the use of database management software for the Z-100 such as CONDOR or DBASE-II, both of which are available through Air Force contracts. Functional Area 3 is best accomplished through the use of spreadsheet or statistical software that can be programmed to perform prescribed analyses via menu driven requests.

The flowchart in Figure 1 shows the flow of information through each of the functional areas described above, along with the type of software considered appropriate to the task. One reason that current spreadsheet or statistical software is not appropriate for the entire system is that it requires the entire database to be loaded into random access memory (RAM), or else develop a complicated scheme of partitioning the database. On the other hand, current database software can directly access the hard or floppy disk storage medium. A disadvantage of the later is that disk access slows the response time significantly for certain types of requests. However, the response times that would be incurred by most functions of the fleet analysis system would be well within those required by our functional specifications.

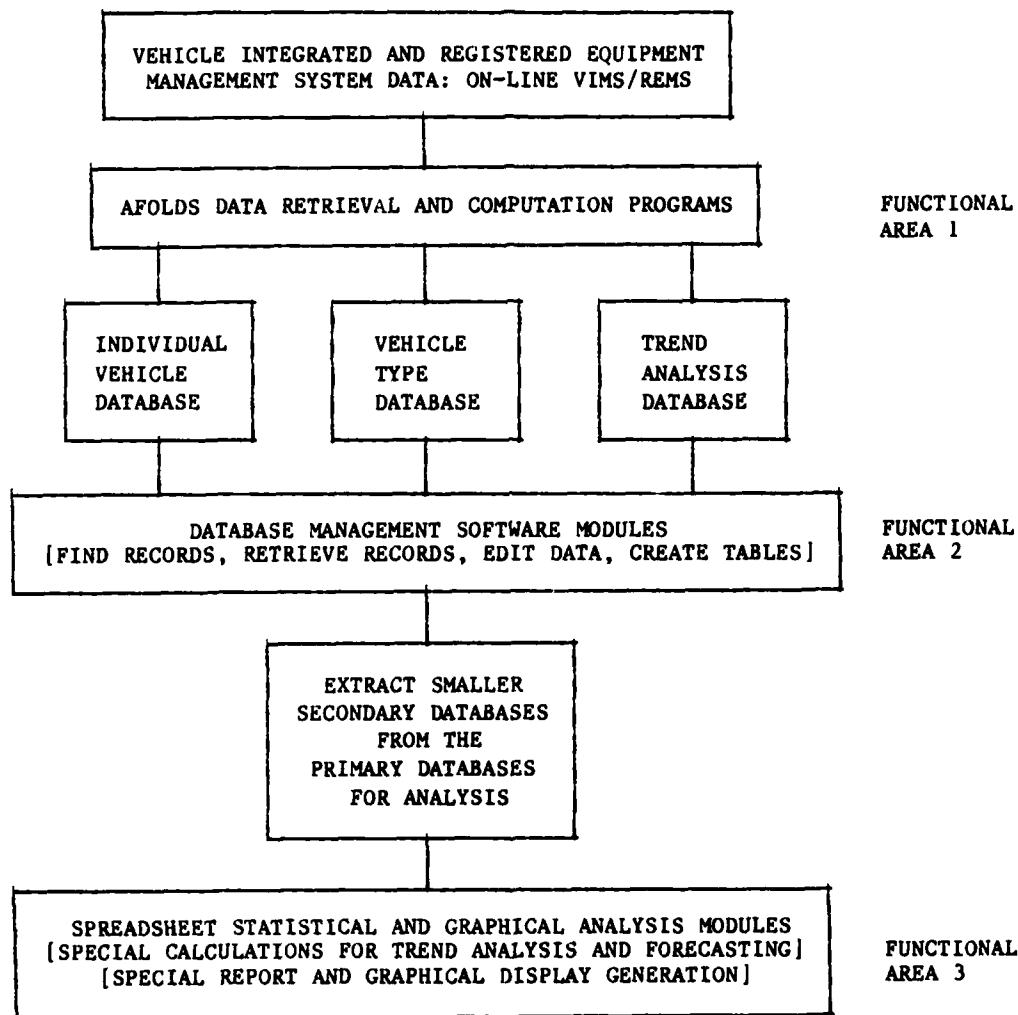


FIGURE 1: FLOW OF INFORMATION THROUGH THE THREE FUNCTIONAL AREAS OF THE VEHICLE FLEET ANALYSIS SYSTEM

## 5. COMPILING A SAMPLE DATABASE

In the initial weeks of research, there were several alternative opinions as to how a database for a system of this type ought to be compiled. An important consideration emphasized at that time was that compiling a small database for system development did not require that the eventual mainframe-to-micro data retrieval procedure be completely developed. Since a suitable method of downloading data from the mainframe had yet to be devised by the Data Systems Design Office (DSDO), we simply compiled our own database by-hand based on VIMS reports for 1983 from Randolph Air Force Base in San Antonio, Texas. In this compilation process, we were careful to include a mix of vehicle types and ages, with some vehicles being retired and others being added to the fleet over the 5 month period for which data was compiled. We also restricted our attention to just a subset of the VIMS/REMS data elements that we identified as being most important to fleet analysis.

Table 1a lists the data elements on individual vehicles that were compiled for months May thru September of 1983 for approximately 30 vehicles. The actual data is not shown in this report because of space constraints. In addition to data on individual vehicles, data was compiled by vehicle type, as shown by Table 2a. Most of this data will be obtained from the REMS portion of the VIMS/REMS consolidated database in the new On-Line Phase IV system. As this information is downloaded to the Z-100 from the Sperry mainframe computer, additional calculations will be performed so as to spare the Z-100 from extensive number crunching computations. Values to be calculated from the raw data are shown in Tables 1b and 2b for both individual vehicles and by vehicle type. Additional calculations can be easily added to the data downloading programs if requested.

## 6. DEVELOPMENT OF A DEMONSTRATION MODULE

Taskings 4 and 5 listed in Section 2 require that the system be subdivided into a series of development modules and that a prototype data display and graphical analysis module be developed and demonstrated. A key reason for first developing the data display module is that it could be programmed and implemented rather quickly on the Z-100 computer to demonstrate potential system capabilities at the Vehicle Fleet Analysis Conference held at AFLMC in July. This strategy proved successful in that it convinced many attendees

TABLE 1a: RAW DATA ELEMENTS COMPILED FOR INDIVIDUAL VEHICLES

1. Management Code (MGT CODE)
2. Registration Number
3. Vehicle Type/Make
4. National Stock Number
5. Owning Command (OWN COM)
6. Using Command (USE COM)
7. Assigned Organization (ASG ORG)
8. Date of Assignment (DATE ASGNMT)
9. Fuel Consumption This Period (F-U T-PER)
10. Use This Period (M/K/H T-PER)
11. Cumulative Use (CUM M/K/H "Miles/Kilometers/Hours")
12. Operations Cost This Period (OPS-C T-PER)
13. Replacement Code (REPL CODE)
14. Maintenance Cost This Period (VDP T-PER)
15. Veh. Downtime for Maintenance This Period (VDM T-PER)
16. Veh. Downtime for Parts This Period (VDP T-PER)
17. Mission Essential Indicator (MISN ESNL)
18. Item Code (ITEM CODE)
19. Use Code (USE CODE)
20. Allowance Source Code (ALL SOR CODE)
21. Base of Planned Use (BASE PLUS)
22. Alternate Storage Location (ALT STR LOC)

TABLE 1b: ADDITIONAL VALUES CALCULATED FOR INDIVIDUAL VEHICLES

1. Operations Cost Per M/K/H This Period
2. Maintenance Cost Per M/K/H This Period
3. Total Cost Per M/K/H This Period
4. Total Vehicle Downtime This Period (VDM+VDP)
5. Percent Vehicle Downtime This Period (VDM+VDP)
6. Hours of Vehicle Use This Period
7. Ratio of Use This Period to the Base Standard
8. Flag Indicator for Use This Period Outside of a Specified Range of the Base Standard
9. Ratio of Use This Period to the Air Force Standard
10. Flag Indicator for Use This Period Outside of a Specified Range of the Air Force Standard

TABLE 2a: RAW DATA ELEMENTS COMPILED BY VEHICLE TYPE

1. Management Code (MGT CODE)
2. Vehicle Description (type, capacity, drive, fuel)
3. Prime National Stock Number (PRIME NSN)
4. M/K/H Code (i.e., how usage is measured)
5. Authorized Quantity
6. Due-In Quantity

TABLE 2b: ADDITIONAL VALUES CALCULATED BY VEHICLE TYPE

1. Number of On-Hand
2. Number of Mission Essential
3. Ratio of Number On-Hand to Number Authorized
4. Ratio of Mission Essential to Number On-Hand
5. Average Age (in years)
6. Average Cumulative Use (M/K/H)
7. Average Use This Period
8. Base Standard Use Per Period
9. Air Force Standard Use Per Period
10. Average Operations Cost This Period
11. Average Maintenance Cost This Period
12. Average Total Cost This Period
13. Average Operations Cost Per M/K/H This Period
14. Average Maintenance Cost Per M/K/H This Period
15. Average Total Cost Per M/K/H This Period
16. Average VDM Per Vehicle This Period
17. Average VDP Per Vehicle This Period
18. Average Total Downtime Per Vehicle This Period (VDM+VDP)
19. Average Percentage Downtime This Period (VDM+VDP)

that an automated system could be tailored to fit their needs so long as a consensus is reached as to just what types of analyses can be best utilize. Once the design constraints for a system of this nature were made clear, the conference attendees were able to identify functions of the system that they felt were most important.

The demonstration module was programmed through the use of Lotus 1-2-3, a widely used spreadsheet software package. This particular software contains programming options that allow the developer of a system to quickly generate command menus for each and every procedure. By doing so, the user is spared from learning formal command sequences and instead can concentrate on the objectives of analysis.

Figure 2 shows a tree diagram of the menu branching logic used in the demonstration package. A detailed explanation of the system is not warranted in a report of this type, but a brief rundown of the menu structure illustrates the flexible and logical sequence of user request options. Menu 1 is designed to permit use of on-line help pages or to gain access to the worksheet environment from which data can be displayed, formatted and graphed. Menu 2 is the main program control menu that enables the user to return to the help pages, load new data or initiate data retrieval and formatting options. In each menu, the Quit option returns the user to the previous menu. Thus, typing Quit successively will eventually lead the user back to Menu 1 from where the entire system can be neatly exited.

Criteria/Find in Menu 3 and its branches is to specify search criteria that are used to find, examine and retrieve matching records. Such criteria can be used to extract a portion of database for specific analyses. Table-Create in Menu 3 is similar to Criteria/Find in that it allows the user to specify precisely which columns of data are to be operated upon in later analyses or formatted into tables. Print and Graph in Menu 3 are the output control options that enable the user to conduct graphical analyses or to format tables for a hardcopy printer. Menu branches from the Print and Graph options allow the user to title and scale tables and graphs.

Tables and graphs can be viewed at the screen, saved to a hard or floppy disk for later use, or printed immediately. A large part of the Vehicle Fleet Analysis Conference was spent debating the most appropriate styles of tabular

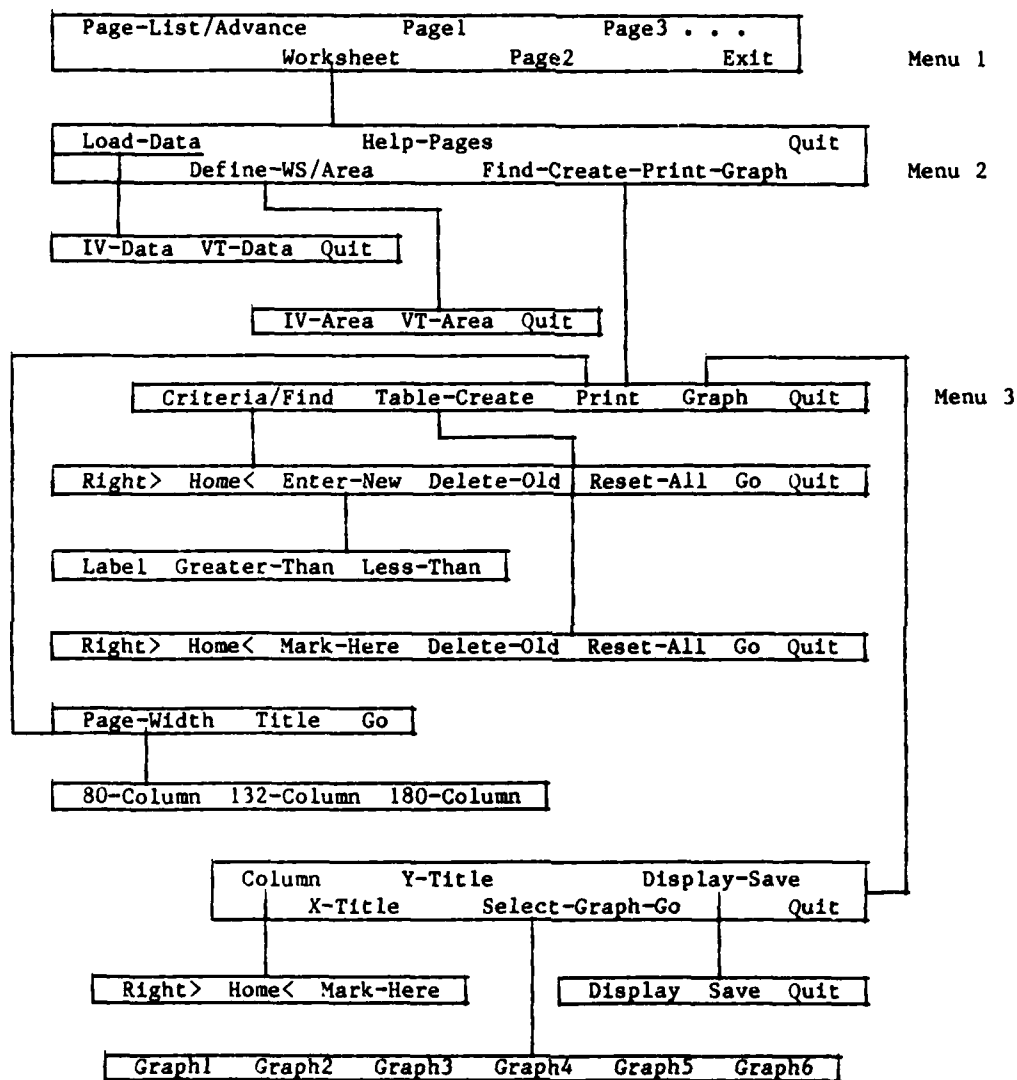


FIGURE 2: TREE DIAGRAM OF MENUS IN THE VEHICLE FLEET ANALYSIS DEMONSTRATION MODULE

and graphic display for different parts of the analysis system, since this is recognized to have a significant impact on the extent to which such output is used for decision making. [7]

## 7. CONCLUSIONS

The conclusions of this research can be categorized under two headings -- those pertaining to lessons learned concerning system development approaches, and those pertaining to the precise configuration of software needed to get the job done. With regard to alternative system development approaches, the speed with which we were able to implement a demonstration module was primarily due to a clear understanding of what our objective was from the start. Our objective was not to develop a final product ready for field testing, since this depended on the input and cooperation of test base personnel. Instead, our objective was to demonstrate potential capabilities of a system so that eventual users could lend their guidance during later stages of system development. This 10-week research project, along with the Vehicle Fleet Analysis Conference, was successful in meeting these objectives.

With regard to the precise configuration of software needed to develop a complete system, we believe that the functional specifications given in Section 3, and the system development plan with the functional areas outlined in Section 4, point correctly towards the best solution. Many technical details must be worked out concerning mainframe-micro communication, as well as the compatibility of database structures used by software programmed for each of the three functional areas. Recommendations concerning these issues are presented next.

## 8. RECOMMENDATIONS

Recommendations for further research and development of a vehicle fleet analysis system are many. Fortunately, we are at the infant stage of development where these recommendations can be addressed before a great deal of human resources are spent in a misguided effort. The most critical issues still to resolve before diving headlong into a full-blown programming effort are:

- a. What are the technical constraints of the mainframe to micro communication link? Can it be implemented directly by the user at the



micro-computer? How rapidly does the transfer of data take place? Can most all extensive computations be completed at the time of downloading? What is the best programming language to perform this task?

- b. How will the current transition from batch-mode VIMS/REMS to On-Line VIMS/REMS impact how and what type of data are retrievable from the mainframe computer? Will the switch in mainframe computers be an advantage or disadvantage to the data retrieval process?
- c. What is the best existing software to handle the database management and analysis capabilities of the system? What new hardware/software breakthroughs are expected in the near future that will impact current technology and permit the system to adopt a significantly different approach?
- d. Just what types of analyses do vehicle fleet managers believe to be the most useful? Is research needed to devise more basic or theoretically correct analyses? Are current methods of reporting or estimating vehicle conditions and costs adequate?

With regard to this last question, I recommend that new research be initiated to evaluate the economic validity of methods used by the Air Force to rate the conditions of their vehicles. At present, various factors of age, recent repair costs and accumulated use are factored into a A to V rating, where any vehicle below J is a candidate for replacement. In this research, we have concentrated more on the framework through which information will be passed to the user, and not on the quality of the information being passed. It can be said, however, that even the best information goes unnoticed if not presented properly. In a paper by Paul Slovic [8] that resulted from research supported by the Office of Naval Research, he writes, "It has become evident that a key element in decision making is the ability to interpret and integrate information items, the reliability and validity of which are imperfect."

A final area of recommended research that is a vital part of the vehicle fleet analysis system, but which was not looked at in this study concerns vehicle dispatching. In the words of Coyle and Bardi [9], "While it may be hard to believe that the daily hectic pace of dispatching can be done by a computer, it is quite likely that we will see new heuristic programs developed to assist the dispatching area."

#### Acknowledgements

The author would like to thank the Air Force Air Training Command, who was the primary sponsor of the CATS-VM project with which this research was coordinated. Co-sponsors of the project were Tactical Air Command and Strategic Air Command. Grateful appreciation is in order for the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education who sponsor the USAF-SCEEE Program that made this research engagement possible.

Conducting research in a supportive environment has a major impact on the quality of the product, and also on the enjoyment of the participant. I am very grateful to the Air Force Logistics Management Center, and especially the Transportation Directorate, for providing a conducive research environment. In particular, I would like to thank Lt. Col. Walter Dzialo, Director of LGT, Major Dan King and Capt. James Van Scotter for their guidance and answers to many detailed questions. Appreciation is also extended to Mrs. Pat Harden for being very helpful and typing various forms, and to Mrs. Ann Bailey for helping to prepare my final report. Their assistance throughout the prototype development process and their comments concerning this final report are appreciated. The content and opinions of this report remain, however, the sole responsibility of the author.

#### REFERENCES

1. Department of the Air Force (1983) "Motor Vehicle Management", Air Force Manual 77-310, Volumes I and II.
2. Ward P. (1984) "The Fleet Maintenance Management Cycle", Transitions, Winter/Spring 1984.
3. Air Force Logistics Management Center (1984) "Computer Assisted Transportation Systems - Vehicle Management: Project Plan", Project No. LT840101, Gunter Air Force Station, AL.
4. Ives B. and Olson M.H. (1984) "User Involvement and MIS Success: A Review of Research", Management Science, Vol. 30(5).
5. Orr K. (1983) "User-Focused System Development Needs Planning and Design", Future Magazine.
6. Rand A.J. (1984) "Producing Viable MIS Reports", Infosystems, June 1984.
7. Remus W. (1984) "An Empirical Investigation of the Impact of Graphical and Tabular Data Presentations on Decision Making", Management Science, Vol. 30, No. 5.
8. Slovic P. (1976) "Towards Understanding and Improving Decisions", Science, Technology and the Modern Navy, Office of Naval Research Report No. 37.
9. Coyle J.J. and Bardi E.J. (1984) The Management of Business Logistics (Third Edition), West Publishing Company.

# AFLMC will review 'bare base' kits

by Capt. Carol A. Walker  
Air Force Logistics Management Center

Picture yourself as a young airman on mobility. When the "balloon goes up," your unit will deploy. But, to where and under what conditions? Just exactly what will be there for you? What kind of facilities will protect you and provide housing, feeding and a place to work? How will you launch the aircraft needed to fight and sustain the mission for which you have been mobilized?

As long as your deployment location has at least a "bare base," that is, a usable runway, taxiway, aircraft parking area, and a source of water that can be purified for drinking, the Air Force can provide the other "mission essential" resources. Those include mobile facilities, utilities, and support equipment that can help turn a bare base into an operational air base almost overnight.

The bare base is part of today's readiness concept — to

deploy forces with shelters and support facilities capable of supporting and launching sustained combat operations with the same independence as bases already in the "theater." Bare base assets provide the minimum support needed for survival of the fighting forces.

**THERE ARE TWO KINDS OF BARE BASE "KITS."** One is made up of mainly soft-walled shelters and support equipment generally used for deployments of short duration. One complete package provides items to bed down a force of over 1,000 people. The other "kit" is of hard-walled construction and contains vehicles, general aircraft maintenance and a "broad base" of logistics support for a sustained operation of a 4,500 person wing. The larger grouping is normally used for longer deployments.

Regardless of the type, billions of dollars will be spent to increase bare base assets. But, how well will these new assets meet the needs of the Air Force? How well can they be maintained? What problems exist that need to be

resolved before new items are purchased? Those are questions the Air Force Logistics Management Center at Gunter will help answer at the direction of Headquarters Air Force.

**THE CENTER WILL DIRECT TEAMS** from other major commands and agencies in gathering data to evaluate bare base assets to be used in a coming major exercise. The teams will look at a bare base package that will support over 1,000 people. They will also evaluate a new kitchen that will be tested. Much of the equipment is newly purchased.

Over a three-month period, from the initial planning stages for a bare base to the time the troops "tear down" and put the package back together, the teams will see how well they work.

The project, said officials, is just one way the AFLMC works to improve logistics support and our "war-winning" capability.

## Conference helps to better vehicle management

**THE FIRST STEP TOWARD** better management of the Air Force's \$3 billion vehicle fleet was taken July 17-18 at the Air Force Logistics Management Center at Gunter.

Those were the dates of the first Air Force-wide Vehicle Fleet Analysis Conference held at Gunter. Center Transportation Project Manager Capt. Jim Van Scotter said the conference was a necessity because major commands and base level vehicle management offices weren't keeping standard records or reports.

In effect, said one official, major commands and the air staff were comparing "apples" and "oranges" when reviewing varied reports on management of the Air Force's enormous fleet.

**CAPTAIN VAN SCOTTER** and Summer Faculty Research Program participant Dr. Bruce Janson have

been developing a management portion of the Computer Assisted Transportation System. The data system, cosponsored by three other major commands, puts a variety of reports and files on vehicles into a standardized computer system.

During the conference a host of major command representatives and people from the air staff discussed how the Air Force's vehicle fleet should be managed and what information was actually needed for decision making. They also discussed how they could best standardize information.

The move toward more standardization and automation was hailed by many. "This is one of the most important efforts in transportation during the last 10 years," said Lt. Col. Thomas West, Headquarters Air Force.

**THE SYSTEM, WHICH WILL BE** prototyped, is

the spring and "fielded" shortly thereafter, is expected to give base level vehicle managers a "standard" data base so they can assess their own abilities to meet day-to-day, exercise and other special requirements that are part of their mission, officials said. AFLMC officials said the move to standardize and automate is just a beginning. They envision a day when a vehicle operations and management information system is ready for Air Force "transporters" worldwide.

"We have a \$3 billion fleet of vehicles and we want to update management by breaking the 'old mold' and bringing microcomputers to base-level vehicle managers," said the AFLMC's Director of Transportation Lt. Col. Walt Dzialis.

Reported by Capt. Carol Walker, Air Force Logistics Management Center.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PHOTOGRAPHIC EMULSIONS FOR PREPARATION OF HOLOGRAPHIC FILTERS

Prepared by:	Dr. Charles R. Jones
Academic Rank:	Associate Professor
Department and University:	Department of Physics North Carolina Central University
Research Location: Beam	Elgin Air Force Base, AFATL/DLMI
USAF Research	Dr. Steve Butler
Date:	September 10, 1984
Contract No:	F49620-82-C-0035

PHOTOGRAPHIC EMULSIONS  
FOR PREPARATION OF  
HOLOGRAPHIC FILTERS

by

Charles R. Jones

ABSTRACT

High resolution photographic emulsions are widely employed for the production of holographic filters to be used in coherent optical processors. Experimental efforts to accurately characterize such emulsions are reported here. Techniques for obtaining transmittance versus exposure are described and some results reported. An analytical procedure for treating some effects of emulsion non-linearities by use of a polynomial representation of the response curve is also presented.

### Acknowledgements

I gratefully acknowledge the support of the Air Force Office of Scientific Research of the Air Force Systems Command which made this summer research activity possible. I am particularly grateful to Dr. Samuel Lambert for his excellent organization of this program at AFATL and to Steve Butler and Dale Fink for the pleasure of working with them during this period.

## I. INTRODUCTION:

Since the advent of lasers, the ability of coherent optical systems to perform certain mathematical operations, e.g. Fourier transformations, correlation and convolution, has been widely recognized and studied. Such an optical processor can achieve very high throughput in the task of pattern recognition. Figure 1 summarizes the basic operation of a typical coherent optical processor. an input scene is recorded on transparency,  $T(x, y)$ , and is placed in the spatial plane ( $P_1$ ) where it is illuminated by a coherent wavefront. Lens  $L_1$  produces the Fourier transform of this input in the frequency plane ( $P_2$ ). Various filters may be inserted into this plane to modify the frequency spectrum before the inverse Fourier transform is produced in the output plane ( $P_3$ ) by Lens  $L_2$ .

Communication theory shows that the optimum linear filter for detection of a target is stationary white noise is a matched filter, i.e. one whose transmission is the conjugate of the target's Fourier transform.<sup>1</sup> When such a filter is placed in the frequency plane, a correlation "spike" will occur in the output plane at the location of the target. This provides a very powerful pattern recognition tool since all of the input transparency, which may contain very large quantity of data, is processed simultaneously and the output appears in plane  $P_3$  at literally the speed of light. However, in many applications the sensitivity of such a processor to the scale and orientation of the target is a severe limitation. For example, a filter which is matched to a target in a particular orientation will not be matched to that target if its orientation is changed by more than a few degrees.

In recent years, much attention has been given to devising methods for achieving scale and orientation invariance. Two of the most promising methods are the use of Mellin-Fourier transforms as suggested by Casasent<sup>2</sup> and the Synthetic Discriminant Function (SDF) approach proposed by Caulfield. Each of these methods involves a preliminary processing step which can be most readily implemented by digital techniques, followed by an optical processor to achieve



high throughput. Thus the most powerful system to accomplish the general task of pattern recognition appears to be a hybrid digital/optical processor which capitalizes on the strengths of each technique.

The Electro-Optical Terminal Guidance Branch of the Air Force Armament Laboratory (AFATL/DLMI) is actively researching the potential of such hybrid systems. This effort involves use of the Image Processing Laboratory (IPL), which has a sophisticated capability for digital image manipulation, and of the Laser Laboratory where a state-of-the-art optical processor is under development.

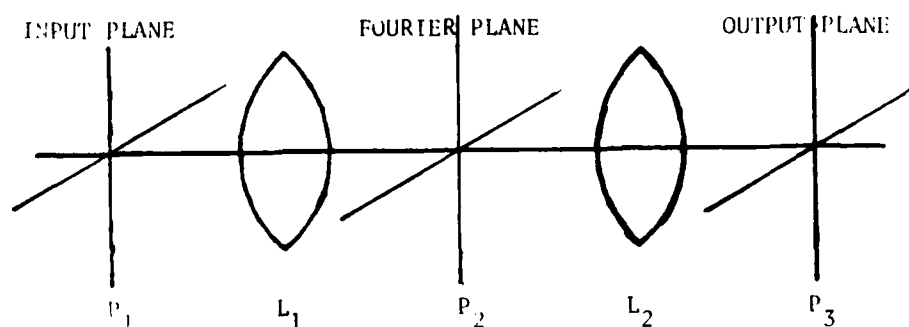


Figure 1 - Typical Coherent Optical Processor

## II. OBJECTIVES:

The objectives which became primary during my stay at AFATL revolved around the critical role which photographic emulsions play in optical processing. These emulsions are presently the highest quality material available for the preparation of transparencies and filters. However, careful storage, processing and calibration are necessary if reproducibly high-quality is to be achieved. Even with such careful quality control, the non-linearity and noise which is inherent in such emulsions must be considered if optimum results are to be achieved.

Two specific objectives were primary:

- (1) To devise a convenient and reliable procedure for the calibration of emulsions.
- (2) To analyze the impact which emulsion non-linearities and noise have upon a coherent optical processor, and to seek those experimental conditions which would minimize this impact.

### III. EMULSION CALIBRATION EXPERIMENTS

#### A. Initial Studies With Laser Source

When used in a coherent optical processor, the essential parameter of any optical element is its amplitude transmittance,  $t(x,y)$ . For a photographic emulsion, this in turn is determined by the energy,  $E(x,y)$ , to which it has been exposed. Thus an emulsion is best characterized experimentally by the transmittance versus exposure ( $t$ - $E$ ) response curve.

To obtain such a curve, the central portion of a 3mW He-Ne laser beam was filtered and expanded to uniformly cover a 4x5 inch photographic plate. The intensity of this beam was measured with an NRC radiometer. A 21-step calibrated photographic step tablet was then placed in the beam just before the photographic plate. Using the known intensity of the original beam and the known transmission of each step on the tablet, the intensity at each point on the photographic plate was determined. The resulting exposure energy in Joules at each point was computed as the product of the intensity (in Watts) and the exposure time (seconds). The exposure time was chosen to give a transmittance of 0.5 near the middle of the exposure range so the  $t$ - $E$  curve could be obtained over a wide range.

When these studies were initially undertaken, three types of high-resolution emulsions were on hand, Agfa 10E75, Agfa 8E75, and Kodak 120. Trial exposures were made on sample plates of each type. Standard development consisted of 5 minutes in Kodak D-19 at approximately 74° (agitation at one minute intervals), followed by one minute in Kodak Stop Bath, four minutes in Kodak Rapid Fix, a 30 minute wash, and one minute in Kodak Photo-Flo prior to drying.

After processing, the intensity transmission of each exposure step was measured with the direct beam of the 3mW laser as a source and the NRC radiometer as the detector. The amplitude transmittance was then calculated as the square root of the intensity transmission. The  $t$ - $E$  curves obtained for the two Agfa emulsions are shown in Figures 2a and 2b. As expected, 8E75 is a much less sensitive emulsion, requiring approximately 50 times as much exposure to produce the same

transmission level as the 10E75. Very significant however, is the response of these emulsions as the exposure approaches zero. The maximum transmittance obtained with the 8E75 is about 95% while with the 10E75 it reached only about 75%. Examination of the receiving records revealed that the 10E75 plates had been on hand in excess of a year while the 8E75 plates were less than 6 months old. Discussions with the manufacturer confirmed that the high level of "base fog" on the 10E75 plates was attributable to aging effects to which the faster plates are quite susceptible. The Kodak 120 plates were several years old and showed such high level of base fog that no attempt was made to measure the t-E curve. These results illustrate the importance of proper storage and careful emulsion characterization.

#### B. Densitometer and Sensitometer Studies

Although the techniques described above for obtaining t-E curves yields satisfactory data, they are time consuming and require the use of laser apparatus. Commercial equipment is available for making test exposures and reading emulsion transmissions. Hence one goal of this effort was to establish simplified calibration procedures using such equipment.

A densitometer (Macbeth TD 504) was made available for use by the Photo-Lab and the previously studied 8E75 plate was measured with this instrument. Such equipment is not ideal for this purpose because it gives data in photographic density,  $D$ , which is related to transmittance by  $t = 10^{-D/2}$ . Nevertheless, with the transmittance calculated from the densitometer data, the resulting t-E curve is in satisfactory agreement with that obtained using the laser/radiometer combination. This is shown in Figure 3.

Investigations were also begun on the use of an EGG sensitometer to perform the calibration exposures. Several complications arise when attempting to replace the laser source by the flashlamp of the sensitometer. First, the emulsion's response to the spectrally broad energy from the flashlamp will be different from its response at the single wavelength of the laser. Second, neither the total exposure energy nor

the spectral distribution of that energy were known for the sensitometer output. It was anticipated that a filter would be required to restrict the sensitometer spectral output to a region near the laser wavelength before reliable calibrations could be performed. Since no suitable filter was immediately available, a preliminary experiment was undertaken. The NRC radiometer, which is calibrated at the laser wavelength, was used to estimate the energy content of a pulse from the sensitometer which was set to  $10^{-4}$  seconds. The result obtained was 13 microJoules/cm<sup>2</sup>. This is, of course, a very approximate value. The calibrated step tablet was then placed on the cover glass of the sensitometer and a 10E75 plate exposed. The exposure energy at each step was estimated from the known transmission of the step tablet using the value obtained from the radiometer. After processing, the amplitude transmittance of this plate was determined and a t-E curve plotted. Figure 4 shows a comparison between this data and the t-E curve previously obtained using the laser source. The degree of agreement is unexpectedly good, however this result should be used with some caution since there may be offsetting error factors which accidentally lead to this agreement. Further studies on this point were postponed pending receipt of a new supply of the 10E75 plates. One conclusion which can be drawn is that the sensitometer energy output is quite sufficient to expose even the very slow emulsions. The  $10^{-4}$  second setting which was used provides the lowest energy and higher settings can be used if needed to offset filter losses.

#### C. Studies on New 10E75 Plates

Near the conclusion of my research period, a new supply of 10E75 plates was received and I immediately began to evaluate these plates. First, an exposure was made using the sensitometer as described above. This plate was processed and its transmittance values determined. The resulting t-E curve is shown in Figure 5. Comparable data is also shown for the old 10E75 plates on the same graph. These two sets of data are consistent with an aging process which decreases the transmission of the old plates at low exposures. Note, that the two sets of data are very similar in their approach to saturation at high exposures.

As a further check, a calibration exposure was made using the laser source as described in Section III A. The t-E data obtained from this plate is also shown on Figure 5. These results are strikingly different from any previous t-E curves, whether obtained by sensitometer or laser exposure. In particular, the sensitivity is much less than specified by the manufacturer. It was difficult to accept the validity of this data so a number of checks were made. The method of exposure, the developer used, the radiometer accuracy and the shutter speed were all examined but no satisfactory explanation could be found. A second set of sensitometer and laser exposures was made and reproduced the same phenomenon. Unfortunately my research period ended before this question could be resolved.

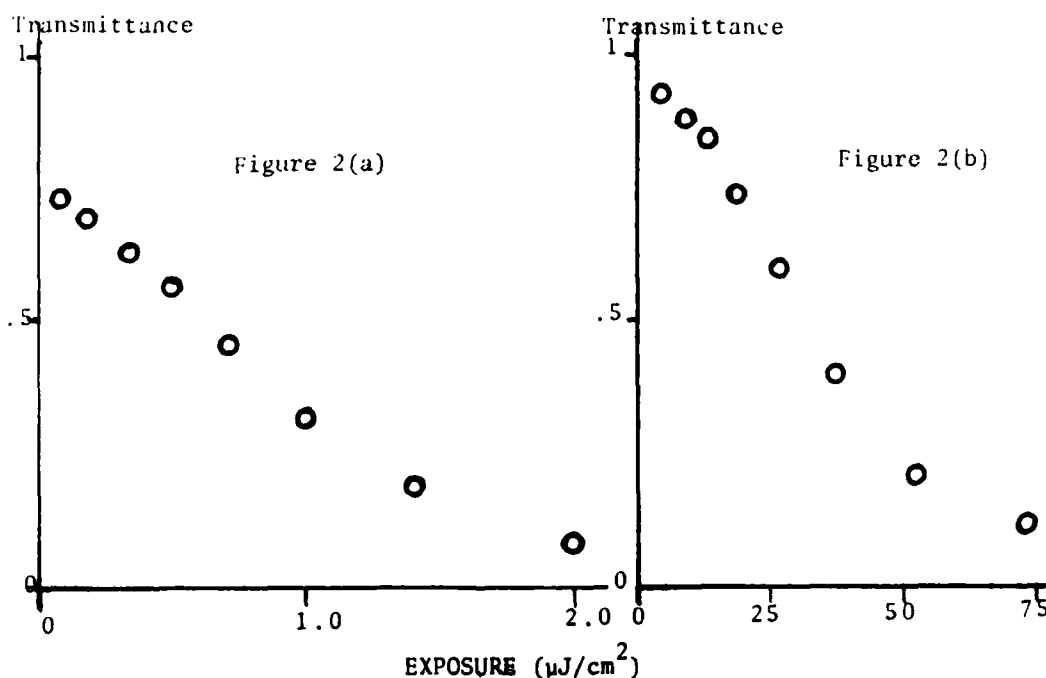


Figure 2 - Transmission versus Exposure for Agfa 10E75 (a) and Agfa 8E75 (b)

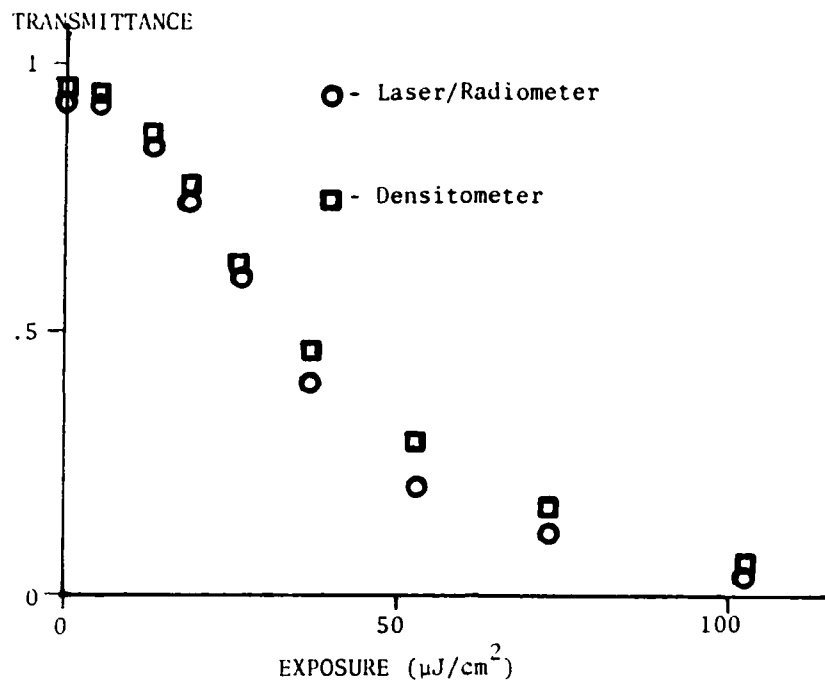


Figure 3 - Comparison of Densitometer Results with Laser/Radiometer Results

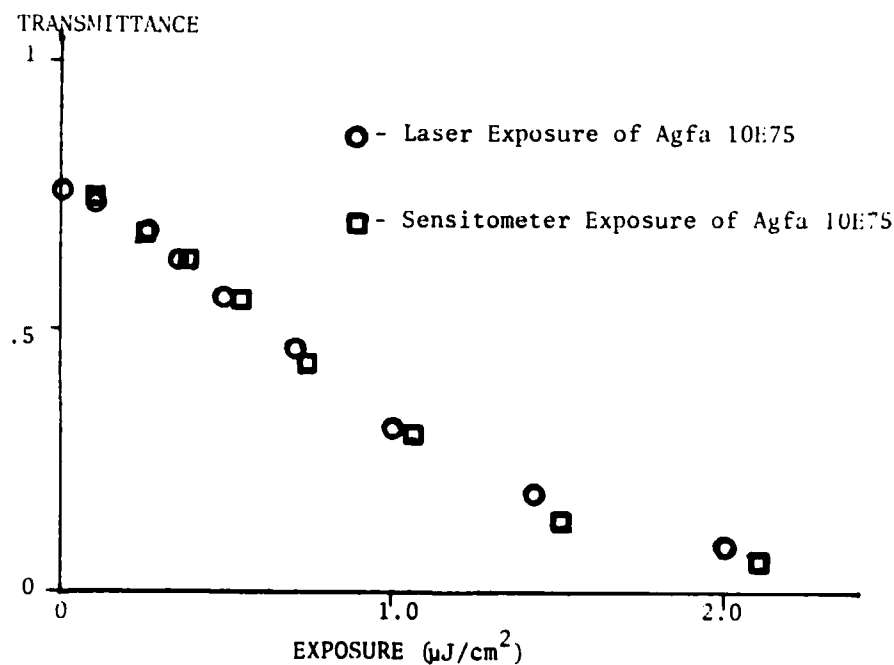


Figure 4 - Comparison of Laser Exposure and Sensitometer Exposure

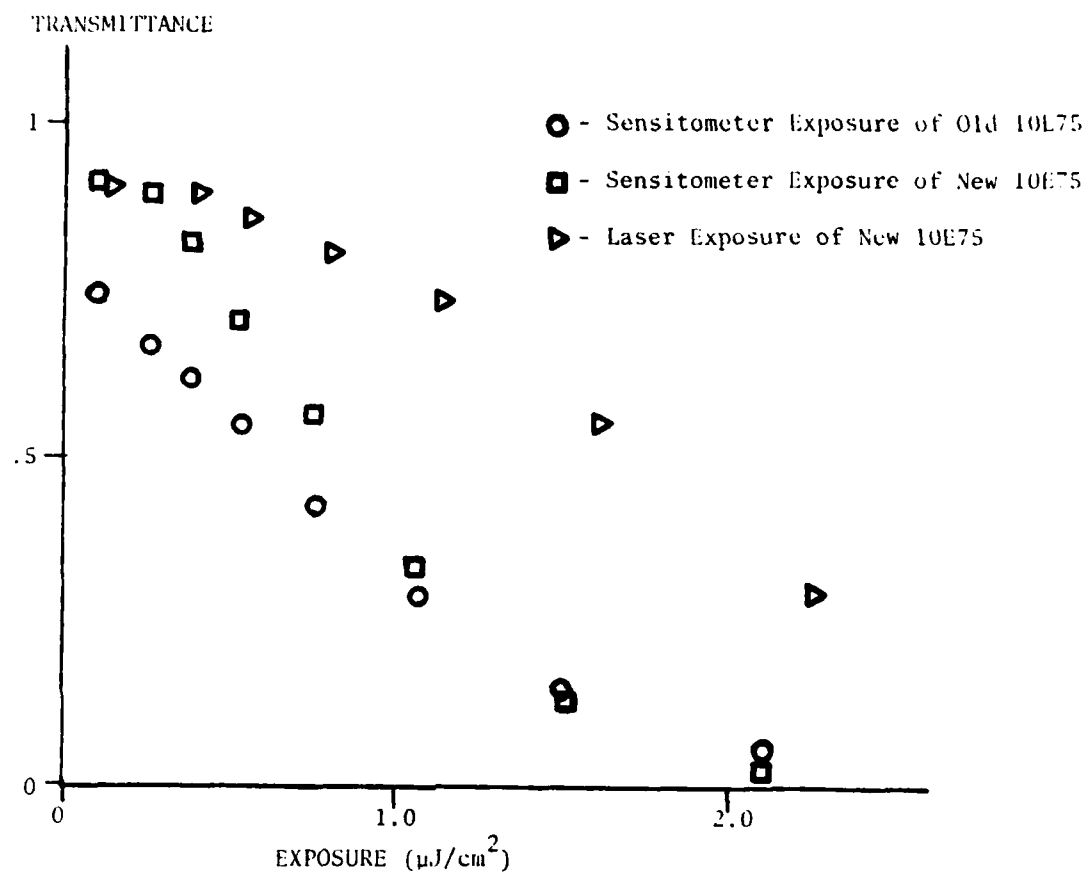


Figure 5 - Comparison of Old and New Agfa 10E75



#### IV. PHOTOGRAPHIC RECORDING OF HOLOGRAPHIC MATCHED FILTERS

##### A. Ideal Case

As noted in the introduction, the optimal filter for detection of a target has transmittance which is the conjugate of the target's Fourier transform. This requires a complex transmittance, which can not be directly implemented on a photographic emulsion. Instead, the emulsion is used as a diffractive element with the phase and amplitude information encoded onto a high-frequency amplitude grating. A typical system for recording such a filter is shown in Figure 6. Lense  $L_1$  forms the Fourier transform of the target,  $S(p) = s(p)e^{i\phi(p)}$ , in plane  $P_2$  and a uniform reference beam  $R(p) = re^{i\alpha p}$  is introduced at an angle  $\theta$  to the optical axis. Here  $p$  is the spatial variable in the Fourier plane (only one dimension is shown for simplicity) and  $\alpha = k \sin(\theta)$ .

A recording media placed in plane  $P_2$  will thus receive an exposure

$$E(p) = |S(p) + R(p)|^2 \tau \quad (1)$$

where  $\tau$  is the exposure time.

Substituting for  $S(p)$  and  $R(p)$ , we obtain

$$\begin{aligned} E(p) &= [\lambda^2 + \lambda^2(p)] \tau + \lambda \tau \lambda(p) [e^{i\phi(p)} e^{-i\alpha p} + e^{-i\phi(p)} e^{i\alpha p}] \\ &= \bar{E} + E_m \end{aligned} \quad (2)$$

where  $\bar{E} = r^2 + s^2(p)$  is the local average exposure and

$$E_m = \lambda \tau \lambda(p) [e^{i\phi(p)} e^{-i\alpha p} + e^{-i\phi(p)} e^{i\alpha p}] \quad (3)$$

is the modulation.

The ideal recording media will have transmittance directly proportional to exposure,  $t = cE$ , for all values of  $E$ . Hence the above exposure would yield

$$T(p) = c \tau [\lambda^2 + \lambda^2(p)] + c \lambda \tau \lambda(p) [e^{i\phi(p)} e^{-i\alpha p} + e^{-i\phi(p)} e^{i\alpha p}] \quad (3)$$

The first term is real and will give rise to wavefronts propagating near the optical axis. The factors  $e^{\pm i\alpha p}$  are linear phase shifts which will give rise to diffracted wavefronts propagating at angles  $\pm\theta$  to the optical axis. Thus, if the recording angle is properly chosen, each of these wavefronts will be separated from the other and from those

which propagate on axis. We note that the second of these terms is proportional to the conjugate of the target's Fourier transform and thus has the desired optimal filter characteristic.

Unfortunately, photographic emulsions do not have the ideal response described above. To be more realistic, one must seek an appropriate model to describe the response of the emulsion to an exposure and to predict the resulting transmittance of the developed emulsion. Such a model can be based upon the experimental response curve, measured as described in Section III. In order to optimize experimental conditions, one wishes to predict the amplitude of the optimal filter term and how it depends upon the various experimental parameters, such as the average exposure and the ratio of reference to signal exposures. A complete model would also describe the origin and behaviour of all noise contributions arising from the emulsion. A search of the literature shows that several authors have treated some aspect of this topic however no completely exhaustive treatment has been found to date.

#### B. Effects of Emulsion Non-Linearities on Amplitude of Filter Term

##### (1) Model based on Hurter-Driffield curve

In normal photographic applications, the response of an emulsion is described in terms of the Hurter-Driffield (H-D) curve which is plot of photographic density,  $D = \log(I_0/I)$ , versus  $\log(E)$ .<sup>3</sup> An idealized form of this curve would be described by

$$D = \begin{cases} D_s & ; E > E_s \\ \gamma [\log(E) - \log(E_{bf})] + D_{bf} & ; E_{bf} \leq E \leq E_s \\ D_{bf} & ; E < E_{bf} \end{cases} \quad (4)$$

where  $s$  denotes saturation,  $bf$  denotes base fog, and  $\gamma$  is the slope of the linear portion of the curve. This model predicts that a linear transmittance versus exposure can only be achieved by making a positive print and developing to a net  $\gamma$  of  $-2$ . If this is accomplished the transmittance becomes

$$T = \begin{cases} T_{bf} & ; E > E_s \\ c(E - E_{bf}) + T_{bf} & ; E_{bf} \leq E \leq E_s \\ T_s & ; E < E_{bf} \end{cases} \quad (5)$$

It should be emphasized here that the transmittances are those of the positive print while the exposures refer to the original negative. With the exposure given by equation 2, this model predicts that the maximum signal amplitude without non-linearity will be obtained when the signal and reference amplitudes satisfy

$$|r - s_{\max}|^2 \tau = E_{bf}, \quad |r + s_{\max}|^2 = E_s \quad (6)$$

where  $s_{\max}$  is the maximum signal amplitude in the Fourier plane. This result can also be stated in terms of the experimental parameters,  $\bar{E}_{\max} = (r^2 + s_{\max}^2) \tau$  and  $K = r^2 / s_{\max}^2$ . These are the maximum average exposure and the reference to signal beam intensities, respectively.

The values predicted for maximum signal without non-linearity are

$$E_{\max} = 1/2 (E_s + E_{bf}), \quad K = \left[ \frac{E_s^{1/2} + E_{bf}^{1/2}}{E_s^{1/2} - E_{bf}^{1/2}} \right]^2 \quad (7)$$

Under these conditions, the maximum amplitude of the diffracted signal is given by  $1/4(t_s - t_{bf})$ . For all values of  $s(p)$  less than  $s_{\max}$  the diffracted signal will be proportional to  $s(p)$ , as desired.

This is the earliest model to be utilized and is included here for completeness, but it is seldom used for discussing amplitude holograms. This is true for several reasons. The transmittance,  $t$ , not the photographic density,  $D$ , is the fundamental parameter in holographic filter performance. The corresponding  $t$ - $E$  curve can be measured directly as easily as it can be inferred from the  $D$ - $\log(E)$  curve. Perhaps more importantly, this model requires the assumption that the filter is a positive print made from the original negative. Few, if any experimentors make use of such a print since it is not essential for filter operation. Lastly, this model assumes that the response is perfectly linear until the limiting exposure values are reached and thus makes no predictions about non-linear effects at lower levels.

## (2) Polynomial approximation of the $t$ - $E$ curve.

The  $t$ - $E$  response curve of an emulsion can be approximated by a linear function only over a very limited range of exposures. A number

of authors have discussed various effects of this non-linear characteristic by employing a polynomial representation of the experimental curve.<sup>4,5,6</sup> A polynomial of degree three is the most common choice, perhaps because it is the lowest degree which can reproduce the obvious inflection point near the middle of the experimental curve. Such a polynomial model provides a convenient closed-form, however it must be used with caution since there is only a limited range of exposures for which it will adequately describe the actual response curve.

As an example, the data shown in Figure 2(b) for Agfa 8E75 was well-fit on the interval 0 to 75 by a cubic polynomial of the form

$$t = C_0 + C_1 E + C_2 E^2 + C_3 E^3 \quad (8)$$

with coefficient values:  $C_0 = 0.947$ ,  $C_1 = -5.07 \times 10^{-4}$ ,  $C_2 = -7.04 \times 10^{-4}$  and  $C_3 = 8.67 \times 10^{-6}$ . It is noteworthy that any cubic polynomial such as this will have an inflection point at an exposure value  $E' = -C_2/3C_3$ . At this point, the slope is maximum and its rate of change is zero, i.e. this point corresponds to the center of the "quasi-linear" region. For the example used here this point occurs at  $E = 27 \mu\text{J}/\text{cm}^2$  and  $t = 0.59$ . Thus it should not be assumed that the most linear region is centered at  $t = 0.5$ .

We can briefly outline here a procedure by which the effects of the emulsion's non-linear response can be predicted. The exposure which is used for filter preparation has the general form given by equation 2 and will result in a transmission given by

$$t = C_0 + C_1(\bar{E} + E_m) + C_2(\bar{E} + E_m)^2 + C_3(\bar{E} + E_m)^3 \quad (9)$$

When like powers of  $E_m$  are collected, this can be written in the form

$$t = A_0 + A_1 E_m + A_2 E_m^2 + A_3 E_m^3 \quad (10)$$

where the A's depend only upon the average exposure  $\bar{E}$ . Substituting for  $E_m$  from equation 3, we can finally express the transmittance as

$$t = t_0 + t_1 + t_2 + t_3 \quad (11)$$

where  $t_0$  is a real quantity,  $t_1$  contains all terms which contribute

to the first diffraction order,  $t_1$  and  $t_2$  contain only terms which appear in the second and third diffracted orders. Thus  $t_1$  can be spatially separated from all other contributions and contains the filter term which is desired.

Writing  $t_1$  out explicitly we obtain

$$t = [(C_1 + 2C_2E_r + 3C_3E_r^2) + (2C_2E_s + 9C_3E_rE_s + 3C_3E_s^2)]E_m \quad (12)$$

where the reference and signal exposures are given by  $E_r = r^2\tau$  and  $E_s = s^2\tau$ . Note that we have collected terms in this expression so that the first term in brackets does not depend upon the signal strength. This term multiplying  $E_m$  will produce a linear reproduction as in the ideal case. The second term in brackets contains all contributions which depend upon the signal strength and hence produce non-linear behaviour. If the first term is designated as the signal coefficient and the second as the non-linear noise coefficient, we can evaluate both the signal amplitude which will be obtained as well as the signal to noise ratio for any choice of reference and signal exposures.

It is clear by inspection that the signal coefficient can be made dominant by choosing  $E_r$  to be large. However, it is easily verified that the signal,  $E_m$ , is maximized when  $E_r = E_s$ . A value for  $E_r$  which is optimum in all respects cannot be immediately identified. The choice will depend upon factors such as the laser power available and the noise level which can be tolerated.

One case is particularly interesting. If we choose  $E_r = E_s$  to maximize  $E_m$  and also choose  $E_r + E_s = E'$ , the value at the inflection point, then one can solve for  $E_r$  and  $E_s$ . The result is  $E_r = E_s = -C_2/6C_3$ . It can be verified by direct substitution that these values make the noise coefficient exactly zero! However, this fortunate condition holds only for this particular signal value and not for all values. If  $E_r$  is fixed at this value while  $E_s$  is allowed to vary, the noise coefficient will increase for  $E_s$  greater or less than  $E_r$ . In particular, for  $E_s$  less than  $E_r$ , it can be shown that the noise will peak at

$E_s = E_r/2$  at which point the signal to noise ratio is 6.2. If this low a signal to noise ratio is tolerable then the above choice will result in maximum signal amplitude and may be the preferred choice.

The expression given in equation 12 is quite general for any t-E curve which can be adequately described by a cubic polynomial. It can be used to explore the effect of varying reference and signal exposures and to predict the behaviour of both the absolute signal amplitude as well as the signal to noise due to non-linearity.

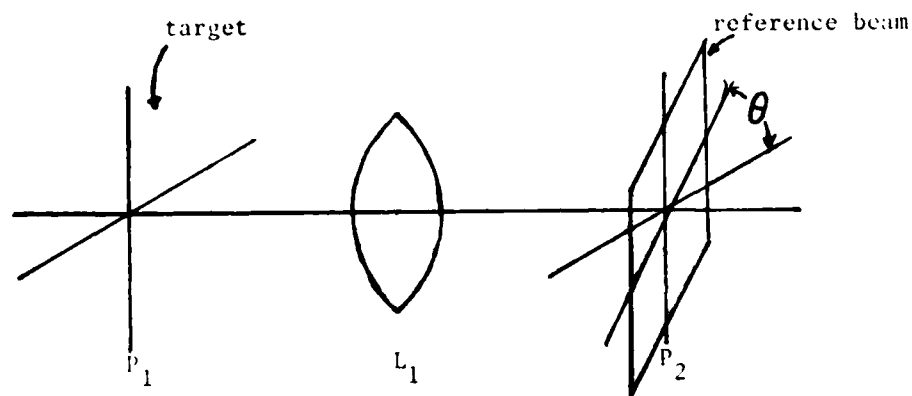


Figure 6 - System for Preparation of Holographic Matched Filters

## V. RECOMMENDATIONS

If reproducible results are to be obtained with photographic emulsions, great care must be given to the details of storage and processing. It is recommended that all plates be kept in refrigerated storage unless they will be used within one month. A means for maintaining a more constant temperature during processing and the use of distilled water for all chemicals is also advisable.

Further studies need to be done in the area of emulsion characterization. An appropriate filter should be obtained for the sensitometer and studies conducted to verify the equivalence of sensitometer and laser exposures. Once this has been accomplished, a sensitometer/densitometer combination will make it possible to routinely calibrate emulsions as needed. I would recommend that one plate be calibrated each time a new box is opened and that careful checks be made on any plates which are kept longer than six months.

A number of follow-on studies could be done to establish experimentally the effects of emulsion non-linearities and whether they are satisfactorily predicted by the use of a polynomial approximation as discussed in Section IV.

#### REFERENCES

1. Davenport, W. B., Jr. and Root, W. L., An Introduction to the Theory of Random Signals and Noise, N.Y., N.Y., McGraw-Hill, 1958.
2. Casasent, D. and Psaltis, D., "Scale Invariant Optical Correlation Using Mellin Transforms," Opt. Comm., V17, 1976, p. 59.
3. Yu, Francis, T. L., Optical Information Processing, N.Y., N.Y., John Wiley and Sons, 1983.
4. Yu, Francis, T. L., "Linear Optimization of Nonlinear Spatial Filters," IEEE Trans. Inform. Theory, IT-17, 1971, p. 524.
5. Lee, Wai-Hon and Greer, Milton O., "Noise Characteristics of Photographic Emulsions Used for Holography," J. Opt. Soc., V61, No. 3, 1971, p. 402.
6. Thomas, Carlton E., "Film Characteristics Pertinent to Coherent Optical Data Processing Systems," Applied Optics, V11, No. 8, 1972, p. 175b.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

FRACTURE BEHAVIOR OF CROSS-PLY GRAPHITE/EPOXY COMPOSITE LAMINATES

Prepared by:	Dr. Walter F. Jones
Academic Rank:	Assistant Professor
Department and University:	Department of Engineering Science and Mechanics The University of Tennessee, Knoxville
Research Location:	Air Force Wright Aeronautical Laboratories/Flight Dynamics Laboratory, Structures and Dynamics Division, Structural Integrity Branch, Fatigue and Fracture Group
USAF Research:	Dr. George P. Sendeckyj
Date:	August 27, 1984
Contract No.:	F49620-82-C-0035

FRACTURE BEHAVIOR OF CROSS-PLY  
GRAPHITE/EPOXY COMPOSITE LAMINATES

by

Walter F. Jones

ABSTRACT

A study of the effect of fiber failure and matrix cracking on the strength of cross-ply graphite/epoxy composite laminates using both analytical and experimental techniques is presented. An improved, second order shear lag theory which models a unidirectional fibrous composite with broken fibers is developed and compared to earlier first order shear lag theories. It is demonstrated that the second order theory provides considerable improvement over first order theories in its predictions of longitudinal fiber stress and transverse matrix stress. Also, an experimental investigation of the internal damage growth in cross-ply graphite/epoxy laminates is described. Acoustic emission techniques are used to monitor damage present at various load levels during static loading. It is shown that the acoustic emission events coming from internal damage such as fiber breaks, matrix cracking, and delamination appear to have very different characteristics in terms of amplitude, duration, rise time, and energy. A discussion of the various characteristics of acoustic emission events and their relation to the actual types of damage present in a laminate is presented.

### Acknowledgement

The author wishes to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the opportunity to spend a very rewarding summer at the Air Force Wright Aeronautical Laboratories, Wright-Patterson AFB, OH. He would like to acknowledge the Flight Dynamics Laboratory, in particular the Fatigue and Fracture Group of the Structural Integrity Branch, for its hospitality and excellent working conditions.

The author wishes to express his deepest appreciation to Dr. George P. Sendekyj for his guidance throughout the summer research period. This research project has been a tremendous learning experience and the author looks forward to working with Dr. Sendekyj on future projects in composite materials research. Finally, the assistance of Mr. Dale Martin of RDP Corporation and Mr. Dennis White of Dunegan Corporation in acquiring state-of-the-art Dunegan System 8000 acoustic emission equipment is gratefully acknowledged.

## I. INTRODUCTION

Fiber-reinforced composite materials provide many advantages to the designer such as high strength, light weight, and the ability to tailor material properties for a given application. Much of the recent research in applied mechanics has been directed toward understanding and characterizing the complex behavior of these composite materials. For example, the designer must know how a composite panel that has some amount of internal damage will perform under loading and how the internal damage will grow until final failure occurs. This knowledge must be realized before future aircraft composite structures can be certified.

Many investigators, including the author, have studied the growth of damage in composite laminates. Some authors [1-8] have studied damage growth from a very fundamental perspective - the unidirectional lamina (all fibers parallel to the load direction). The fracture behavior of unidirectional composites has been modeled by several investigators using mathematical models based on the classical shear lag assumption. These models, which assume that the fibers carry all of the axial load and the matrix transmits only shear between fibers, have in their most recent forms taken into account fiber breakage, matrix yielding, and matrix splitting [4-7]. The author used the shear lag model to predict damage growth in unidirectional boron/aluminum composite laminates [8].

There is clearly a need to understand the failure of more complicated composite laminates. For example, internal damage in the form of matrix splitting in unidirectional graphite/epoxy is controlled by matrix normal stresses as well as matrix shear stress. It has been shown in [9] that most shear lag models predict values for the matrix normal stress that are inconsistent with elasticity theory. Thus

improved shear lag theories such as [9] must be developed to give accurate approximations for the matrix normal stress as well as the matrix shear stress and longitudinal fiber stress. More complicated stacking sequences, such as cross-ply laminates, must be studied to determine the effect of internal damage on laminate strength. The cross-ply laminate (a laminate having fibers oriented parallel and perpendicular to the load direction) seems to represent the next logical step in analyzing more complex laminates. It should be possible to model the cross-ply laminate with broken fibers and matrix cracks using a correct shear lag theory.

Many investigators have attempted to observe damage in composites in the laboratory using various nondestructive evaluation techniques. One technique that has seen considerable use in recent years is acoustic emission (AE). Using ultrasonic transducers, the AE equipment can detect inaudible but measurable acoustic pulses that occur within a specimen when some type of damage is formed. Several investigators have conducted investigations using acoustic emission [10-14], however, no one has really explained the nature of AE events generated by different types of internal damage such as fiber breakage and matrix cracking. Some have theorized that the events that arise from fiber breakage have very distinct characteristics in terms of amplitude when compared to other events. While it has been shown in [14] that boron fiber breaks can be distinguished from other events, the breakage of much smaller graphite fibers should be much more difficult to separate. In fact, no investigator has clearly described the acoustic emissions generated from graphite/epoxy laminates during loading.

In this investigation, a study of fracture mechanisms in graphite/epoxy composite laminates will be initiated. Accurate shear lag

models for the composite with broken fibers will be developed including the study of second order theories. Cross-ply Gr/E laminates will be studied to determine where and when the initial damage occurs during a static load cycle and how the damage grows until failure of the laminate. Finally, acoustic emission techniques will be used to detect internal damage in the cross-ply laminates during static loading.

## II. OBJECTIVES

The major objective of this project was to begin a study of the effect of matrix cracking on the fracture behavior of cross-ply graphite/epoxy composite laminates. Some preliminary work was required before the actual study of matrix cracking could begin. Improved modeling techniques for the composite with broken fibers needed to be developed and a certain amount of experimental data needed to be generated. Thus the specific objectives of the summer research effort were:

- (1) To develop shear lag theories for the composite with broken fibers that were consistent with the theory of elasticity and gave correct values for longitudinal fiber stress, matrix shear stress, and matrix normal stress.
- (2) To perform a series of tests using cross-ply graphite/epoxy laminates having various numbers of 90° plies (perpendicular to the load direction) while monitoring the growth of internal damage using acoustic emission techniques.
- (3) To study the characteristics (amplitude, rise time, duration, counts, etc.) of the acoustic emissions originating from various types of internal damage such as fiber failure, matrix cracking, and delamination.

### III. SECOND ORDER SHEAR LAG THEORY

All of the shear lag theories in the literature [1-8] provide reasonable values for the longitudinal fiber stress. The various theories are actually very consistent from one to another in the longitudinal fiber stress values predicted. However, only the most recent shear lag theory [9] provides reasonable values for the matrix normal stress. All of these theories use first order finite difference approximations for the derivatives that arise from the equilibrium equations. It has been suggested that higher order approximations should be used to give accurate values for the fiber and matrix stresses.

In order to derive a second order shear lag theory, consider a unidirectional fiber-reinforced composite lamina undergoing plane deformation. On the micro scale, the composite can be modeled as an elastic matrix reinforced by equally spaced, parallel, elastic fibers as shown in Figure 1a. Defining  $b$  and  $h$  as the width and spacing of the fibers, respectively, a typical element of the composite can be isolated as shown in Figure 1b.

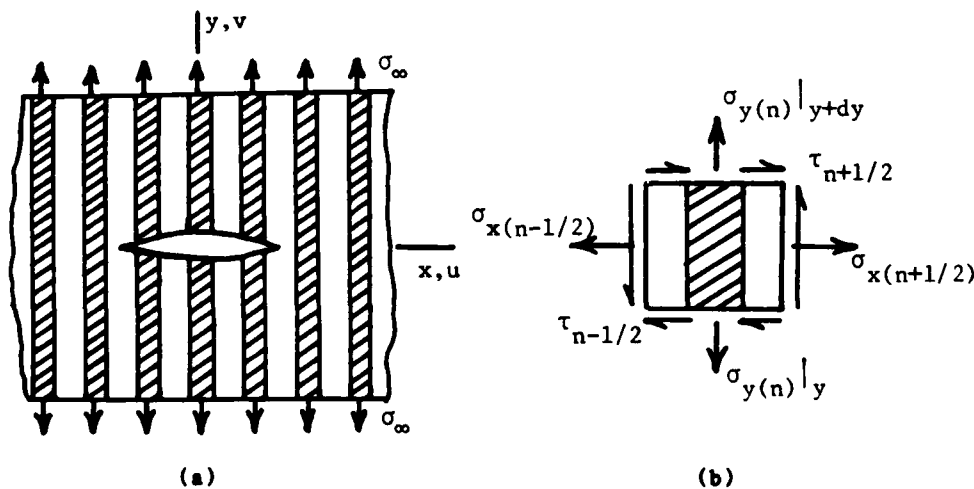


Figure 1. Unidirectional Lamina Geometry and Typical Element

As seen in Figure 1b, the  $n^{\text{th}}$  typical element is bounded by the  $(n+1/2)$  and  $(n-1/2)$  mid-node lines. The equilibrium equations for the element are

$$\begin{aligned}\sigma_{x(n+1/2)} - \sigma_{x(n-1/2)} + h(\tau_{n+1/2,y} + \tau_{n-1/2,y})/2 &= 0 \\ \tau_{n+1/2} - \tau_{n-1/2} + h\sigma_{y(n),y} &= 0\end{aligned}\quad (1)$$

The subscripts following the commas denote partial derivatives with respect to the subscript variable, that is,  $f_{,x} = \partial f / \partial x$ . The quantities in the above equations are merely the appropriate stresses evaluated at the boundaries of the element. As detailed in [15], the constitutive relations in terms of the mid-node quantities are

$$\sigma_{x(n+1/2)} = C_{11}u_{(n+1/2),x} + C_{12}v_{(n+1/2),y} \quad (2a)$$

$$\sigma_{y(n)} = C_{12}u_{n,x} + C_{22}v_{n,y} \quad (2b)$$

$$\tau_{n+1/2} = C_{66}[u_{(n+1/2),y} + v_{(n+1/2),x}] \quad (2c)$$

where  $u_n$  and  $v_n$  are the displacements of the  $n^{\text{th}}$  line and the terms with subscripts  $(n+1/2)$  are displacement quantities at the mid-node lines. The appropriate shear lag expressions can then be determined by writing the mid-node quantities in the following manner:

$$\begin{aligned}f_{n+1/2} &= f_n + (h/2)f_{n,x} \\ f_{n-1/2} &= f_n - (h/2)f_{n,x}\end{aligned}\quad (3)$$

The constitutive relations in terms of nodal quantities can easily be obtained by substituting the finite difference approximations that are consistent with the order of the theory being derived. For example, the second order shear lag theory can now be derived by using the following second order central difference formulas:

$$\begin{aligned}f_{,x} &= (f_{n-2} - 8f_{n-1} + 8f_{n+1} - f_{n+2})/12h \\ f_{,xx} &= (-f_{n-2} + 16f_{n-1} - 30f_n + 16f_{n+1} - f_{n+2})/12h^2\end{aligned}\quad (4)$$

Upon substitution of Equations (2), (3), and (4) into Equations (1) and rearranging terms, the equilibrium equations become



$$\begin{aligned}
& C_{11}(-u_{n-2} + 16u_{n-1} - 30u_n + 16u_{n+1} - u_{n+2}) \\
& + h(C_{12} + C_{66})(v_{n-2,y} - 8v_{n-1,y} + 8v_{n+1,y} - v_{n+2,y}) + 12h^2 C_{66} u_{n,yy} = 0 \\
& C_{66}(-v_{n-2} + 16v_{n-1} - 30v_n + 16v_{n+1} - v_{n+2}) \\
& + h(C_{12} + C_{66})(u_{n-2,y} - 8u_{n-1,y} + 8u_{n+1,y} - u_{n+2,y}) + 12h^2 C_{22} v_{n,yy} = 0
\end{aligned} \tag{5}$$

It should be noted that these equilibrium equations are consistent with the second order finite difference approximations of the governing equilibrium equations in terms of displacements for an orthotropic elastic material. Thus, the new second order shear lag theory is consistent with elasticity and should provide reasonable fiber stresses and transverse matrix stresses.

Equations (5) can be normalized by letting

$$\begin{aligned}
\eta &= y/h, \quad U_n = u_n/h, \quad V_n = v_n/h \\
F &= C_{22}/C_{66}, \quad E = C_{11}/C_{66}, \quad L = C_{12}/C_{66}, \quad H = 1 + L
\end{aligned} \tag{6}$$

The resulting differential-difference equations can then be reduced to ordinary differential equations by assuming that the normalized displacements  $U_n$  and  $V_n$  are the Fourier coefficients of functions  $\bar{U}(\eta, \theta)$  and  $\bar{V}(\eta, \theta)$ . For example,  $U_n$  can be written as

$$U_n = (1/2\pi) \int_{-\pi}^{\pi} \bar{U}(\eta, \theta) e^{in\theta} d\theta \tag{7}$$

with the inverse formula

$$\bar{U}(\eta, \theta) = \sum_{n=-\infty}^{\infty} U_n e^{-in\theta} \tag{8}$$

Substitution of the transformed expressions for the normalized displacement quantities into Equations (5) gives the following equations to be satisfied:

$$\begin{aligned}
3\bar{U}_{,\eta\eta} - E(1-\cos\theta)(7-\cos\theta)\bar{U} + iH(\sin\theta)(4-\cos\theta)\bar{V}_{,\eta} &= 0 \\
3F\bar{V}_{,\eta\eta} - (1-\cos\theta)(7-\cos\theta)\bar{V} + iH(\sin\theta)(4-\cos\theta)\bar{U}_{,\eta} &= 0
\end{aligned} \tag{9}$$

These equations can be satisfied using techniques described in detail in [15]. It should be noted at this point that Equations (9) are only a small degree more complex than the corresponding first order equations,

given in [9] as:

$$\begin{aligned}\bar{U}_{,\eta\eta} - 2E(1-\cos\theta)\bar{U} + iH(\sin\theta)\bar{V}_{,\eta} &= 0 \\ F\bar{V}_{,\eta\eta} - 2(1-\cos\theta)\bar{V} + iH(\sin\theta)\bar{U}_{,\eta} &= 0\end{aligned}\tag{10}$$

Therefore, the second order theory provides increased accuracy, does not substantially increase the computational time, and can easily be used to work the problem of a unidirectional composite lamina with broken fibers.

Upon definition of appropriate stress and displacement boundary conditions, the broken fibers problem can be solved by numerically evaluating the integrals that arise from the transform and then solving the set of simultaneous algebraic equations for the Fourier coefficients. Once the Fourier coefficients are determined, the appropriate expressions for stresses and displacements at various points in the laminate can be evaluated. This procedure is detailed in [15], where the actual values of the integrals are given.

At this point, consider the problem of a composite laminate with one broken fiber and a unit load at infinity. The material constants defined in Equations (6) are taken as  $E = 2$ ,  $F = 28$ , and  $L = 0.6$  which corresponds to a typical graphite/epoxy composite laminate. The resulting stress distribution is given in Figure 2 for the present second order theory. The longitudinal fiber stresses predicted by the second order theory are slightly higher than the first order results given in [9]. However, the transverse matrix stresses are considerably higher in the second order results near the broken fiber. It should be noted that early shear lag theories [1-8] produced transverse matrix stresses that were even further from the correct value. In fact, most of these theories produced matrix stresses that had the wrong sign when compared to elasticity theory.

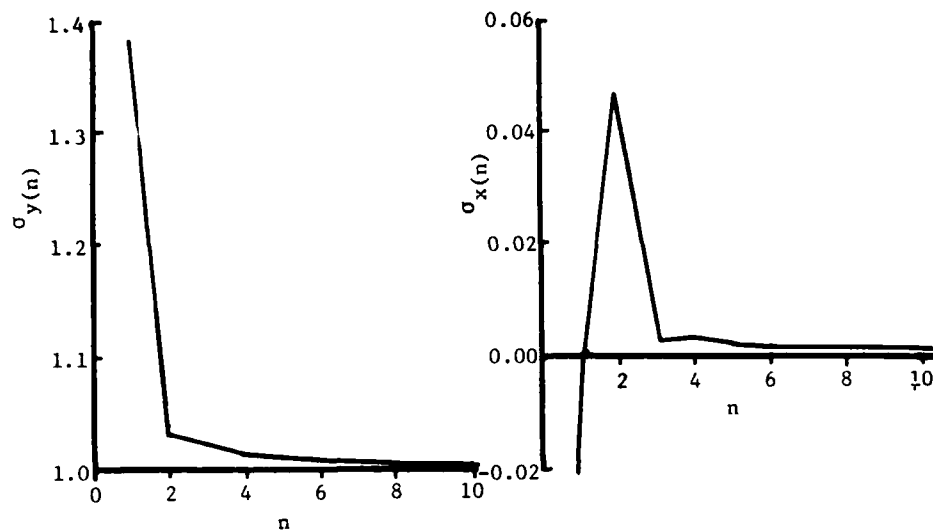


Figure 2. Stress Distribution for Composite with Single Broken Fiber.

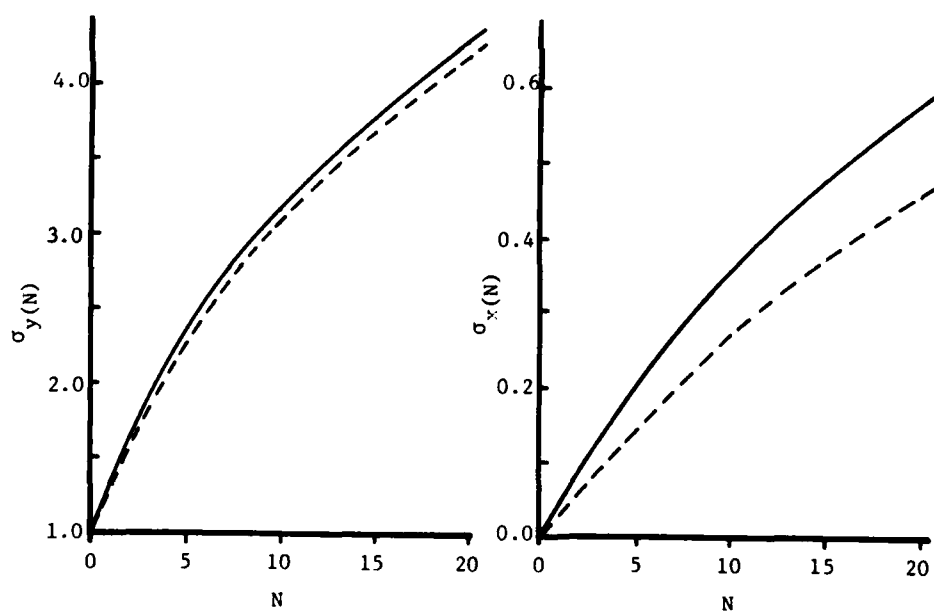


Figure 3. Maximum Stress as a Function of  $N$ : Comparison of the First Order (---) and Second Order (—) Theories.

To emphasize the difference in the theories, Figure 3 shows the maximum longitudinal fiber stress and the maximum transverse matrix stress for different numbers of broken fibers. It is again noted that while little difference exists in the predicted fiber stresses, the matrix stresses for the second order theory are roughly twenty-five percent higher than those predicted by the first order theory. It is thus clear that the higher order theory does a much better job of modeling the stress singularity that exists at the end of the broken fibers.

#### IV. ACOUSTIC EMISSION STUDY OF CROSS-PLY GRAPHITE/EPOXY LAMINATES

The second major objective of this study was to attempt to monitor the growth of damage in cross-ply graphite/epoxy composite laminates during static loading. Acoustic emission (AE) techniques were to be used to give the investigators some insight as to when and where the internal damage formed. Examples of this internal damage are fiber breakage, matrix cracking, and delamination. It had been suggested that the AE events generated by various types of damage would have different characteristics such as amplitude, rise time, and duration.

A series of cross-ply AS4/3502 graphite/epoxy laminates were manufactured having various numbers of 0° and 90° plies. These laminates were cut into specimens that were approximately 48.25 mm (1.9 in) wide. Aluminum tabs were bonded on the ends of the specimens since they are much less prone to generate acoustic emission signals during loading than plastic tabs. Roughly 40 specimens were tested during the summer program using a 35-kip servohydraulic testing machine. It should be noted that several unidirectional specimens were tested to provide strength data for the AS4/3502 graphite/epoxy. It was determined that

the strength of the material was approximately 2100 MPa (303 ksi).

Several acoustic emission systems were used during the study, ranging from an early single-channel Dunegan/Endevco 3000 to the state-of-the-art Dunegan Corporation System 8000. Several tests were performed using a Dunegan 920 Locator with a Model 922 External Memory Unit. This system allows the investigator to monitor only the AE events that occur within a specified location range and to determine the amplitude distribution of these events. The Dunegan 8000 also provides the capability to isolate only the events occurring within a specified location range. However, the 8000 system has analysis capabilities that are several orders of magnitude greater than those of the earlier systems. The 8000 can determine the amplitude, rise time, duration, estimated energy, and numerous other characteristics of the acoustic emissions during the test and can then generate plots of these parameters versus other parameters such as elapsed time or location. In other words, this system gives the investigator an almost unlimited capacity to analyze AE signals in order to determine their origins.

The Dunegan Corporation 8000 System was used to monitor the damage progression for a series of cross-ply laminates and to generate various graphs which illustrated the characteristics of the AE events. Two sensors were mounted on each specimen so that linear location of each AE event could be accomplished. The specimens were then loaded to failure in the testing machine. After the series of tests was completed, the 8000 was used to generate plots of events, counts, amplitude, rise time, duration, and estimated energy (all versus elapsed time) for each specimen tested. Space does not permit the author to show a complete set of graphs, however, typical graphs for a  $[90_2/0_2/90]_S$  specimen are shown in Figure 4. These graphs depict the total counts and events

versus elapsed time, the rise time for each event versus elapsed time, and the number of events at each amplitude. Only the counts and events could be monitored using earlier systems, indicating the vast improvement in analysis capability of the 8000 system. A much more complete set of graphs for a series of tests can be found in [16].

The Dunegan 8000 also allows the investigator to filter the aforementioned graphs for certain ranges of amplitude, location, and other characteristics. It was noted that many of the events occurred outside the sensors due to grip noise, etc. Thus the same plots as in Figure 4 were generated for only those events occurring within the sensors. Typical graphs with the location filtering are shown in Figure 5. It should be noted that with the location filtering, a large number of events in the 40-50 dB amplitude range are missing. These events were in earlier investigations attributed to matrix cracking. It now appears, however, that the low amplitude events are due primarily to grip noise rather than matrix cracking and that the fiber failures and the matrix cracks have amplitudes that are in the same range. Furthermore, it was discovered in this investigation that the amplitude of a matrix crack event depends on the thickness of the 90° ply. The laminates with single 90° layers had amplitude peaks in the 60-70 dB range while those with two-ply thick 90° plies had peaks in the 85-92 dB range. It has been suggested that fiber failures also have amplitudes in the 60-95 dB range, with boron fiber fracture events occurring at approximately 95 dB. It will therefore be quite difficult to distinguish between fiber and matrix failure on the basis of amplitude. The graphite fiber failures do appear, however, to be low energy events when compared to the matrix cracking events.

It is clear that other acoustic emission event characteristics must

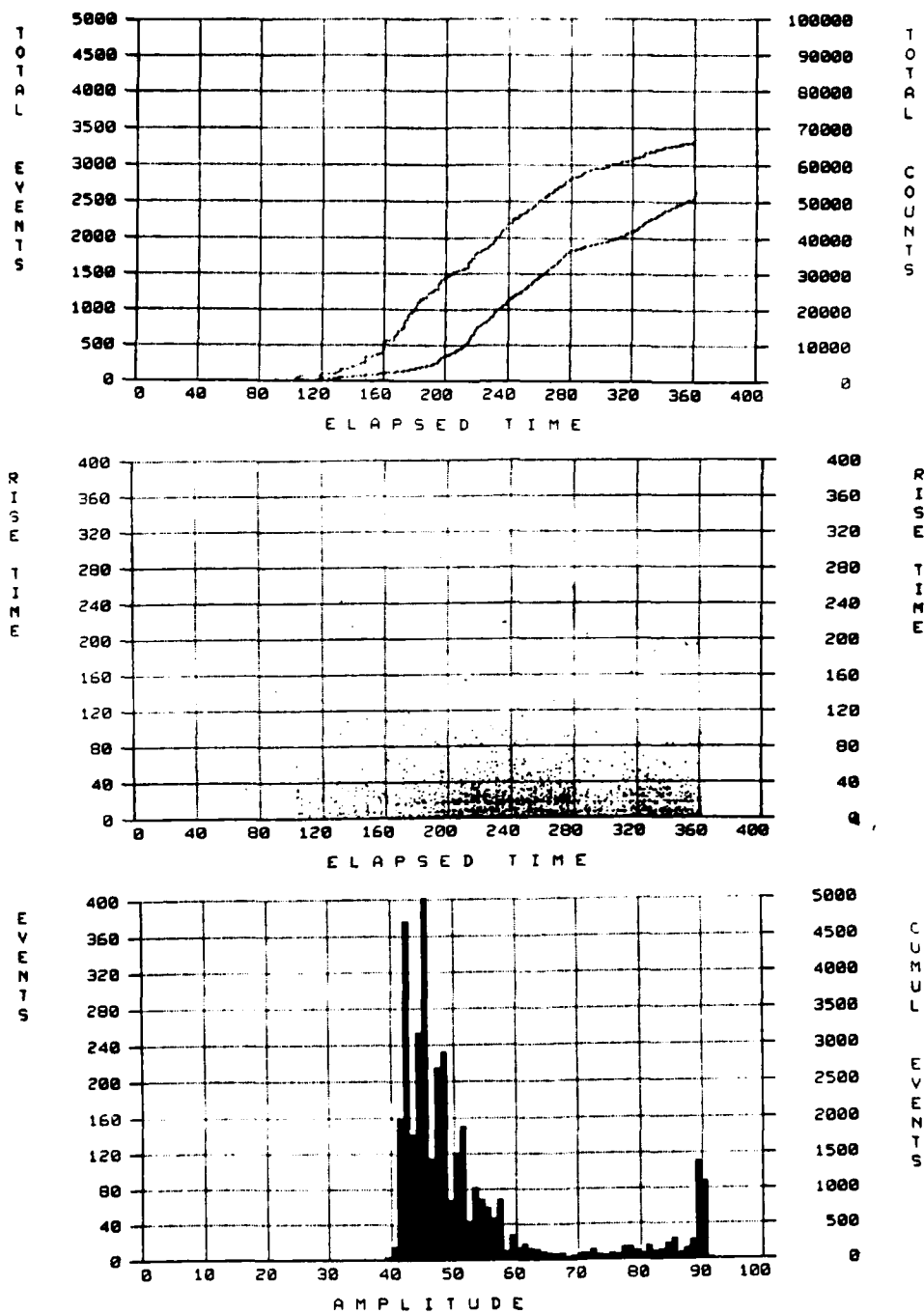


Figure 4. Typical Acoustic Emission Plots for All Recorded Events.

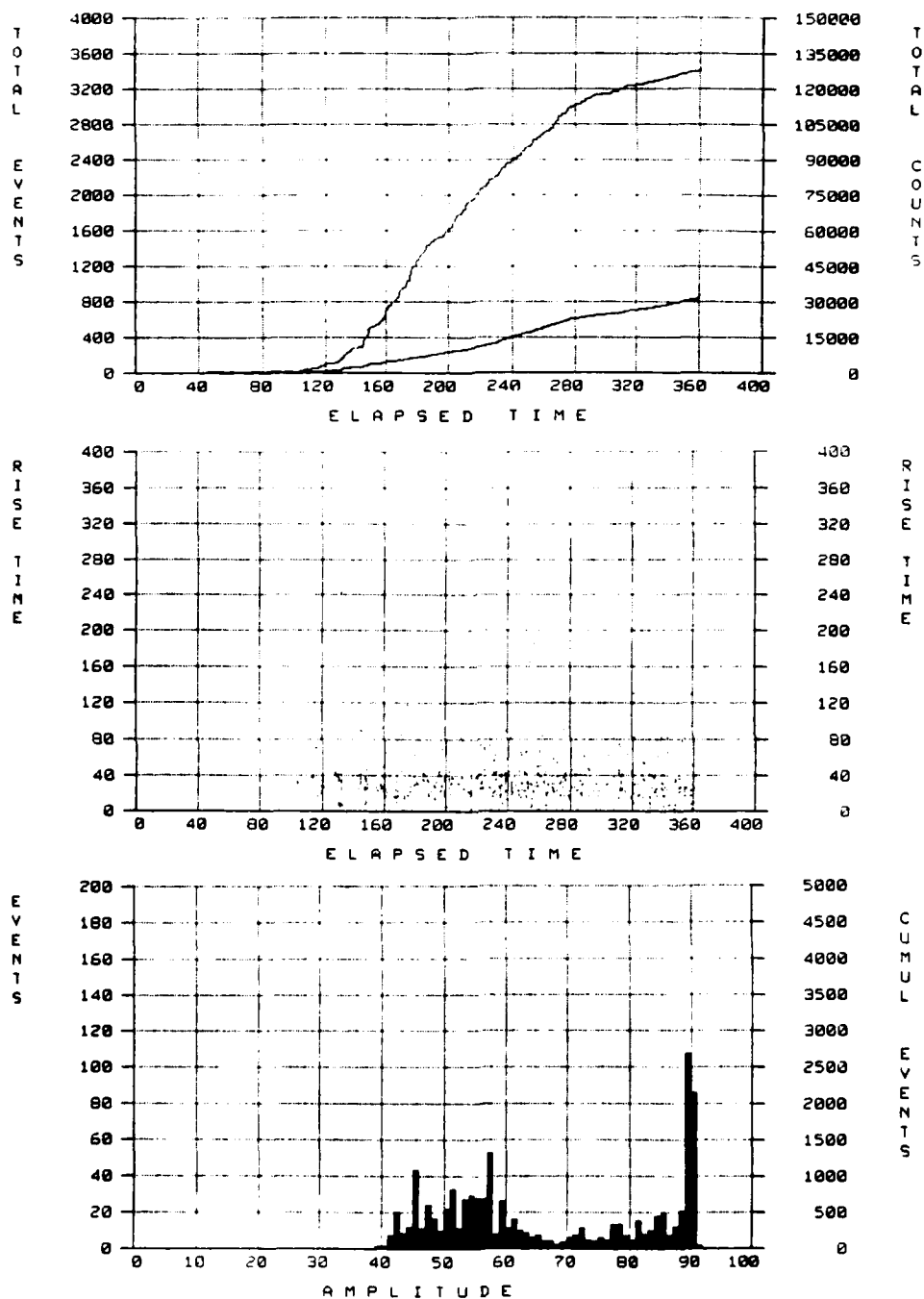


Figure 5. Typical Acoustic Emission Plots for Events Originating between Sensors.



be examined in order to distinguish between different types of damage within a composite. The rise time, duration, amplitude, and estimated energy must all be studied to determine their true significance and any acoustic emission system that is to be useful must be able to eliminate any events coming from outside the desired test section.

#### V. RECOMMENDATIONS

While the results from this study are significant on their own, the the main objective of this summer program from the beginning was to develop both analytical and experimental methods that could later be used to analyze more complex composite laminates. The developed second order shear lag theory is consistent with elasticity theory and provides a considerable increase in accuracy over first order theories without a substantial increase in computational time. The acoustic emission event characteristics have been studied in greater detail than in previous investigations in an effort to determine the relation between acoustic emission data and actual damage growth. However, both of these areas need further investigation to fully realize their significance.

The shear lag theory developed in this study models a composite lamina with broken fibers but with no other damage present. The next step will be to modify the solution to incorporate matrix yielding and splitting. These additions will greatly improve the models capacity to describe the actual failure modes observed in unidirectional composites. The author has modeled matrix yielding and splitting before in [6,7] and used a shear lag theory to study matrix yielding in boron/aluminum composites [8]. The shear lag models used in those studies were capable of giving reasonable estimates of the longitudinal fiber stress, while giving incorrect results for the transverse matrix stress. It is clear

that the new second order theory should be much more capable of modeling the matrix splitting which occurs in graphite/epoxy, a phenomenon thought to be controlled by the matrix normal stress.

After the addition of matrix damage to the new shear lag model, the final step will be to add the capacity to model two adjacent half-planes of different materials. The model can then be used to study cross-ply laminates with cracks in the 90° plies. These cracks are thought to cause significant stress concentrations in the 0° fibers, thereby reducing the strength of the cross-ply laminates. The actual mechanism for this strength reduction and final failure is not clearly understood, and the new modeling techniques developed in this study should provide insight into the problem.

The acoustic emission results from this study have only begun the determination of the true characteristics of AE events generated by various forms of internal damage. Further analysis of acoustic emission signals needs to be accomplished using state-of-the-art equipment as has been done in this investigation. While the author has studied the amplitude, rise time, duration, and other parameters for acoustic emission events, the parameters that actually distinguish a fiber break from a matrix crack have yet to be determined. It does seem that current AE systems, unlike some early systems, have the potential to accomplish this objective upon further study. The author is anxious to use state-of-the-art AE equipment to continue the study after the summer program is over. The end objective of most acoustic emission investigators is to be able to detect and identify internal damage in real structures with no initial defects present. It is only then that acoustic emission will become a useful nondestructive evaluation tool to be used by research organizations and in industry.

#### REFERENCES

1. Hedgepeth, J. M., 'Stress Concentrations for Filamentary Structures,' NASA TN D-882, 1961.
2. Zweben, C., 'An Approximate Method of Analysis for Notched Unidirectional Composites,' Engineering Fracture Mechanics, Vol. 6, 1974, pp. 1-10.
3. Eringen, A. C., and Kim, B. S., 'Stress Concentration in Filamentary Composites with Broken Fibers,' Princeton University Technical Report No. 36, September 1973, ONR Contract N-00014-67-A-0151-0004.
4. Goree, J. G., and Gross, R. S., 'Analysis of a Unidirectional Composite Containing Broken Fibers and Matrix Damage,' Engineering Fracture Mechanics, Vol. 13, 1980, pp. 563-578.
5. Goree, J. G., and Gross, R. S., 'Stresses in a Three-Dimensional Unidirectional Composite Containing Broken Fibers,' Engineering Fracture Mechanics, Vol. 13, 1980, pp. 395-405.
6. Goree, J. G., Dharani, L. R., and Jones, W. F., 'Mathematical Modeling of Damage in Unidirectional Composites,' NASA Contractor Report 3453, August 1981.
7. Dharani, L. R., Jones, W. F., and Goree, J. G., 'Mathematical Modeling of Damage in Unidirectional Composites,' Engineering Fracture Mechanics, Vol. 17, 1983, pp. 555-573.
8. Jones, W. F., and Goree, J. G., 'Fracture Behavior of Unidirectional Boron/Aluminum Composite Laminates,' NASA Contractor Report 3753, December 1983.
9. Sendekyj, G. P., and Jones, W. F., 'A Consistent First-Order Shear Lag Theory for Unidirectional Composites with Broken Fibers,' Engineering Fracture Mechanics. (Submitted for publication.)
10. Rollins, F. R., Jr., 'Acoustic Emission from Boron/Aluminum Composites during Tensile Fracture and Fatigue,' Tech. Report No. 6, Office of Naval Research, October 1971.
11. Grenis, A. F., and Levitt, A. P., 'Acoustic Emission of Graphite and Boron Fiber-Reinforced Aluminum Composites during Deformation and Fracture,' Army Matl. and Mech. Res. Center AMMRC-TR-75-15, June 1975.
12. Carlyle, J. M., 'Imminent Fracture Detection in Graphite/Epoxy Using Acoustic Emission,' Experimental Mechanics, Vol. 18, No. 5, 1973.
13. Hamstad, M. A., 'Acceptance Testing of Graphite/Epoxy Composite Parts with Acoustic Emission,' NDT International, Vol. 15, 1982, pp. 307-314.

14. Awerbuch, J., and Madhukar, M., 'Monitoring Damage Accumulation in Filament-Wound Graphite/Epoxy Laminate Coupons during Fatigue Loading through Acoustic Emission,' Journal of Reinforced Plastics and Composites, Vol. 3, 1984, pp. 2-39.
15. Jones, W. F., and Sendecky, G. P., 'A Second Order Shear Lag Theory for Unidirectional Composites with Broken Fibers,' Journal of Applied Mechanics. (Submitted for publication.)
16. Sendecky, G. P., and Jones, W. F., 'Acoustic Emission Detection of Internal Damage in Graphite/Epoxy Composites,' AFWAL/Flight Dynamics Laboratory Technical Report. (In preparation.)

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE OPTIMAL CONSTRUCTION OF SYNTHETIC DISCRIMINANT FUNCTIONS FOR OPTICAL

MATCHED FILTERS

Prepared by:	Robert R. Kallman
Academic Rank:	Professor
Department and University:	Department of Mathematics North Texas State University Denton, Texas 76203
Research Location: Eglin AFB, Florida 32542	U. S. Air Force Armament Laboratory, Guided Weapons Division, Electro-Optical Terminal Guidance Branch
USAF Research:	Captain James Riggins
Date:	July 25, 1984
Contract No:	F49620-82-C-0035

THE OPTIMAL CONSTRUCTION OF SYNTHETIC DISCRIMINANT FUNCTIONS FOR OPTICAL  
MATCHED FILTERS

by

Robert R. Kallman

ABSTRACT

Synthetic discriminant functions (SDF's) for optical matched filters have potential use for pattern recognition. A variety of large intricate programs were designed, written, debugged, and tested. The purpose of these programs is to compute optimal SDF's modulo certain constraints and compare them to one another. A guiding principle throughout this work is that objective numerical criteria be applied to determine what is a "good" or "optimal" SDF. These programs were tested using 36 infrared tank images. For example, one of the programs picks out the best 4 from the 36 tank images and manufactures an SDF from them. Two other programs, whose purposes are to discretize an SDF in an optimal manner so that an optical hologram of it can actually be manufactured, were designed, written, debugged, and tested. These two programs are somewhat different from one another, for each tries to drive its own special error estimate to 0. There are many, many interesting and important questions which remain to be investigated in this area of research. They are discussed in this report.

### Acknowledgement

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing me with the opportunity to spend a very worthwhile and interesting ten weeks at the Air Force Armament Laboratory, Eglin AFB, Florida. I would like to acknowledge the Air Force Armament Laboratory, in particular the Electro-Optical Terminal Guidance Branch and its Image Processing Laboratory, for its hospitality and excellent working conditions.

Very special thanks go to Captain James Riggins and Steve Butler for introducing the present research topic to me and patiently explaining its many aspects to me. Thanks also go to a great many individuals at the Armament Laboratory for informative scientific and technical discussions and for many kindnesses.

## I. INTRODUCTION.

Imagine a two-dimensional infrared image  $f(x_1, x_2)$  of a scene which contains an object of interest. Consider the following operations on  $f$ . Map  $f(x_1, x_2)$  to its Fourier transform  $F(f)(k_1, k_2)$ , multiply by the Fourier transform of a suitable filter (the synthetic discriminant function or SDF)  $F(h^\#)(k_1, k_2)$ , take the inverse Fourier transform of the product to obtain the convolution  $(f * h^\#)(x_1, x_2)$ , and measure the magnitude  $|f * h^\#|^2(x_1, x_2)$ . Here,  $h^\#(y_1, y_2) = h(-y_1, -y_2)$ , so  $(f * h^\#)(x_1, x_2)$  may be viewed as the inner product of  $f$  and the translate of  $h$  by  $(-x_1, -x_2)$ . The functions  $f$  and  $h$  will be real valued, so no complex conjugates are needed. Actually, some preprocessing can (and should, for a variety of reasons to be discussed later) be done before  $F(f)$  is passed through the filter. For example  $f$  should first be edge enhanced ( $f$  is sent to  $F(f)$ , a suitable disk about the origin is removed from  $F(f)$ , and  $F^{-1}$  of the result is taken), and the then edge enhanced image should be energy normalized ( $f$  is sent to  $f / \|f\|_2$ , where  $\|f\|_2$  is the usual  $L^2$ -norm of  $f$ ) to make the total energy in the scene equal to 1. All of these operations can be carried out almost instantaneously, for Fourier transforms and their inverses can be carried out by lenses, and the important multiplication and filtering step can be carried out by focusing the light wave  $F(f)$  through a lense hologram made from  $F(h^\#)$ . If  $h$  is suitably constructed and scaled, the objects of interest should be centered at the points  $(x_1, x_2)$  such that  $|f * h^\#|^2(x_1, x_2) = 1$ . In reality  $f$  will probably be a 512 x 512 pixel image and  $h$  will be 32 x 32 pixels in size. One can think of the filter as operating by instantaneously placing translates of  $h$  all over  $f$ , taking the corresponding inner products, and indicating those places where the



magnitude of the inner products are large. These should be the places where objects of interest are located. This, in very rough form, is the matched optical filtering process via an SDF.

Needless to say, the correct construction of  $h$  is of the utmost importance if this scheme has any hope of working. An early attempt along these lines was made by Grumman Corporation, who simply took a number of transparencies of the object of interest, at a variety of aspects and angles, and overlaid them.<sup>14</sup> Hence, Grumman started with  $m$  images  $f_1, \dots, f_m$  and let  $h = f_1 + \dots + f_m$ . If  $f_1, \dots, f_m$  are mutually orthogonal, this is a good idea, course in general they are not.

In the past few years a generalization of the Grumman approach was suggested by Casasent et al., who proposed to take a number of infrared pictures  $f_1, \dots, f_m$  of the object of interest and choose  $h$  to be a suitable linear combination of  $f_1, \dots, f_m$ .<sup>4, 5, 6, 10, 11</sup> So in theory  $h = a_1 f_1 + \dots + a_m f_m$  for some suitable choice of constants  $a_1, \dots, a_m$ . Thinking of the  $f_i$  ( $i = 1, \dots, m$ ) as vectors in some high dimensional Euclidean space,  $h$  is expressed as a linear combination of the vectors  $f_1, \dots, f_m$ . To determine the  $a_i$  ( $i = 1, \dots, m$ ), take the inner product of  $h$  with each  $f_j$  and use the bilinearity of the inner product to obtain  $1 = \langle h, f_j \rangle = a_1 \langle f_1, f_j \rangle + \dots + a_m \langle f_m, f_j \rangle$ , where  $\langle \dots \rangle$  denotes the inner product between vectors. If the  $m \times m$  matrix  $(\langle f_i, f_j \rangle)$  is nonsingular (as it most probably is for quite different images  $f_1, \dots, f_m$ ), then there is a unique choice for the  $a_i$  ( $i = 1, \dots, m$ ), and they can be determined easily by solving a system of  $m$  equations in  $m$  unknowns.

For some reason the notion apparently has persisted that one should not use all of the original images  $f_1, \dots, f_m$  to manufacture the SDF  $h$ ,

but instead should use a subset  $p$  ( $< m$ ) of them. Alledged reasons for not using all of the images include potential dynamic range constraints or problems with correlating on clutter.

## II. OBJECTIVES.

The initial part of my SCEEE fellowship was devoted to designing from scratch a variety of programs to generate the best SDF's I could from a training set of images and to compare them to each other. The training set consisted of  $512 \times 512$  tank images, but were 0 outside of rows 200 to 400 inclusive. These images were dirty, in the sense that they did not consist of tanks in a zero background, as would be desired, but were images with a very bright background included. The images were previously edge enhanced and biased. They were furnished to me in the form of a computer tape, each image consisting of a string of  $512^2 = 262144$  integers between 0 and 255, representing intensity levels at each pixel of the image. The guiding principle throughout the initial phase of my work was that notions such as "good" or "best" be determined by concrete numerical criteria. In general, the programs try to drive a least squares error down to 0. The computing was done on a VAX 750/VMS 3.5 in the Image Processing Laboratory.

## III. DESCRIPTION OF THE PROGRAMS USED TO MANUFACTURE SDF'S FROM THE TANK DATA.

Let  $f_1, \dots, f_{36}$  be the tank images, thought of as vectors in a high dimensional ( $512^2 = 262144$ ) Euclidean space. The following concepts and programs, with minor modifications, apply to any number of images, not just 36, and of any size, not just  $512 \times 512$ . In all of the calculations it fortuitously turns out that the only thing one really needs to know about the  $f_i$ 's is the symmetric  $36 \times 36$  matrix  $(\langle f_i, f_j \rangle)$ , which should be

calculated and stored first. As a measure of error made by a potential SDF, I chose the least squares error:  $LSE(h) = |\langle h, f_1 \rangle - 1|^2 + |\langle h, f_2 \rangle - 1|^2 + \dots + |\langle h, f_{36} \rangle - 1|^2$ . This procedure is extremely plausible, for such a measure of error has proven useful over the past 200 years in astronomy and statistics.

The following is a listing of some of the SDF's I calculated, giving their method of calculation and the least squares error for each. They are named and numbered in the order in which I calculated them.

SDF1:  $LSE(SDF1) = 0.0$ . This is the theoretically perfect SDF and is a linear combination of all 36 tank images. So  $SDF1 = a_1 f_1 + a_2 f_2 + \dots + a_{36} f_{36}$ , for some choice of constants  $a_1, \dots, a_{36}$ . The  $a_j$ 's must satisfy the 36 equations  $a_1 \langle f_{36}, f_1 \rangle + a_2 \langle f_2, f_1 \rangle + \dots + a_{36} \langle f_{36}, f_1 \rangle = 1$ , for  $i$  between 1 and 36. They are easily calculated by Gaussian elimination.

SDF2:  $LSE(SDF2) = 0.061511+$ . This SDF is a linear combination of 6 tank images. It was calculated by exhaustively checking all  $(36 \text{ choose } 6) = 1947792$  subsets of the 36 tank images, and for each fixed subset of 6, calculating that linear combination which makes the LSE as small as possible. It is easy to check that  $LSE(h)$  is a convex function of  $h$ , so a local minimum for  $LSE(h)$  is a global minimum for  $LSE(h)$ . If the  $f_i$ 's are independent vectors and  $h$  is restricted to be a linear combination of them, it is easy to check that  $LSE(h)$  becomes uniformly unbounded as  $h$  increases, so a global minimum for  $LSE(h)$  exists. Suppose  $g_1, \dots, g_6$  is a subcollection of 6 out of the 36 tank images. We would like to find numbers  $a_1, \dots, a_6$  so that  $h = a_1 g_1 + \dots + a_6 g_6$  minimizes  $LSE(h)$  over all possible choices of the  $a_i$ 's. The above reasoning indicates that a minimum exists and is assumed when the partial derivate of  $LSE(h)$  with respect to

each  $a_i$  is 0. Doing this for each  $a_i$  gives us 6 linear equations which must be satisfied. They are  $m_{i1}a_1 + \dots + m_{i6}a_6 = b_i$ , where  $m_{ij} = \langle g_i, f_j \rangle = \langle f_j, g_i \rangle + \langle g_i, f_2 \rangle \langle f_2, g_j \rangle + \dots + \langle g_i, f_{36} \rangle \langle f_{36}, g_j \rangle$  and  $b_i = \langle g_i, f_1 \rangle + \langle g_i, f_2 \rangle + \dots + \langle g_i, f_{36} \rangle$ . Given the  $a_i$ 's and  $b_i$ 's as above, a little algebra shows that  $LSE(h) = 36 - a_1b_1 - \dots - a_6b_6$ . So the program to compute SDF2 proceeds as follows: select 6 out of the 36 tank images  $g_1, \dots, g_6$ , find  $a_1, \dots, a_6$  by solving one set of 6 equations in 6 unknowns, and compute  $36 - a_1b_1 - \dots - a_6b_6$ . Select that subset of 6 which makes this last number as small as possible, and manufacture SDF2 from them by computing  $a_1g_1 + \dots + a_6g_6$ . The images selected were 4, 12, 16, 24, 29, and 32.

SDF3:  $LSE(SDF3) = 0.085728+$ . This SDF is a linear combination of 6 tank images. It was calculated in a step-by-step orthogonalization procedure from SDF1. Let  $g_1$  be that tank image so that the orthogonal projection of SDF1 onto the line spanned by  $g_1$  is largest. Take the projection of SDF1 and of  $f_1, \dots, f_{36}$  onto the orthocomplement of  $g_1$  to obtain SDF1' and  $f_1', \dots, f_{36}'$  (only 35 of which are now nonzero), and repeat the process 5 more times. Keep track of the index chosen each time to get the 6 desired original images  $g_1, \dots, g_6$ . The calculations are easy, for  $g_1$  is that image  $f_i$  so that the angle between SDF1 and  $f_i$  is as small as possible. That is,  $g_1$  is that  $f_i$  so that  $|\langle SDF1, f_i \rangle| / (\|SDF1\| \cdot \|f_i\|)$  is as large as possible. Once  $g_1$  is chosen, the iteration is simple, for  $\langle f_i', f_j' \rangle = \langle f_i, f_j \rangle - (\langle f_i, g_1 \rangle \langle f_j, g_1 \rangle / \|g_1\|^2)$  and  $\langle SDF1', f_i' \rangle = \langle SDF1, f_i \rangle - (\langle SDF1, g_1 \rangle \langle f_i, g_1 \rangle / \|g_1\|^2)$ . So repetition is easy. The images selected, in the order in which they were chosen, are 17, 12, 4, 29, 16, and 24. SDF3 is then that linear combination of these images which

minimizes the LSE.

SDF3A:  $LSE(SDF3A) = 0.271316+$ . This SDF is a linear combination of 6 tank images. The choice of the images is done in the same manner as was done for SDF3, and so the images are the same. However, SDF3A was chosen to be the theoretically perfect SDF on these 6 images. It is simple to check that SDF3A is the orthogonal projection of SDF1 onto the subspace spanned by the 6 selected images.

SDF4:  $LSE(SDF4) = 0.085728+$ . This SDF is a linear combination of 6 tank images. To find the images, an exhaustive search of all  $(36 \text{ choose } 6) = 1947792$  subsets of 6 tank images is made. The subset of 6 chosen is that one such that the orthogonal projection of the theoretically perfect SDF1 onto their span is as large as possible. As in SDF3A, this projection is simple to calculate, for it coincides with the theoretically perfect SDF made from the 6 working images. SDF4 is then that linear combination of the images chosen which minimizes the LSE. The images chosen were 4, 12, 16, 17, 24, and 29, the same images chosen by SDF3. This is a fluke (c.f. CSDF3 and CSDF4, to be discussed later).

SDF4A:  $LSE(SDF4A) = 0.271315+$ . SDF4A stands in the same relation to SDF4 as SDF3A does to SDF3.

SDF5:  $LSE(SDF5) = 0.080739+$ . This SDF was calculated in the same manner as was SDF2, except that SDF5 is a linear combination of 5 images. The images chosen were 4, 12, 17, 18, and 29. Notice that the best 5 images are not a subset of the best 6 images.

SDF6:  $LSE(SDF6) = 0.103963+$ . This SDF was calculated in the same manner as was SDF2 and SDF5, except that SDF6 is a linear combination of 4 images. The images chosen were 4, 12, 17, and 29.

SDF7:  $LSE(SDF7) = 0.258666+$ . This SDF is a linear combination of 6 tank images. They are chosen in a step-by-step orthogonalization procedure. Roughly speaking, the first image chosen is that one which contains as much information as possible about all of the other images. The numerical measure for this information is taken to be the sum of the squares of the cosines of the angles between all of the images. So the first image chosen,  $g_1$ , is that image  $f_1$ , so that  $(\langle f_1, f_1 \rangle^2 / \|f_1\|^2 \|f_1\|^2) + \dots + (\langle f_1, f_{36} \rangle^2 / \|f_1\|^2 \|f_{36}\|^2)$  is as large as possible. All vectors are now projected onto the orthocomplement of  $g_1$ , as in the calculation of SDF3, and the process is repeated 5 more times. The images selected, in the order in which they were chosen, are 16, 28, 6, 22, 11, and 36.

SDF7A:  $LSE(SDF7A) = 0.753411+$ . SDF7A stands in the same relation to SDF7 as SDF3A does to SDF3.

SDF8:  $LSE(SDF8) = ?$  This SDF stands in the same relation to SDF7 as SDF4 does to SDF3. This was not run, for I estimate its computation would have taken more than 50 hours of CPU time.

SDF8A:  $LSE(SDF8A) = ?$  SDF8A stands in the same relation to SDF8 as SDF3A does to SDF3. It was not run for the same reasons that SDF8 was not run.

Suppose one chose to make an SDF from the images 1, 2, 5, 16, 19, and 36. The very best SDF that can be made with these images has  $LSE = 0.183932+$ . If one took these same 6 images and took the orthogonal projection of SDF1 onto their span, and used this orthogonal projection as an SDF, it had  $LSE = 0.247391+$ .

Using DeAnza equipment, the tank images were now extracted from the background and placed into 256 x 256 arrays due to concern that the SDF's

might be correlating on the clutter in the background and not on the tanks themselves and because working with 512 x 512 images consumed inordinate amounts of CPU time. I then repeated all of my previous calculations on this new data. The following are the results. In general, CSDF- was manufactured in exactly the same manner as SDF-, except that the cleaned up images were used instead.

CSDF1:  $LSE(CSDF1) = 0.0$ . This SDF is a linear combination of all 36 naked tank images.

CSDF2:  $LSE(CSDF2) = 0.870515+$ . The images used were 14, 16, 24, 29, 32, and 33.

CSDF3:  $LSE(CSDF3) = 1.28928+$ . The images used were 17, 16, 33, 14, 32, and 12.

CSDF3A:  $LSE(CSDF3A) = 2.731533+$ . The images used were 17, 16, 33, 14, 32, and 12.

CSDF4:  $LSE(CSDF4) = 1.141493+$ . The images used were 12, 14, 16, 17, 31, and 33. Note that the images used to manufacture CSDF4 do not coincide with the images used to manufacture CSDF3.

CSDF4A:  $LSE(CSDF4A) = 2.160517+$ . The images used were 12, 14, 16, 17, 31, and 33.

CSDF5:  $LSE(CSDF5) = 1.046511+$ . The images used were 14, 17, 24, 29, and 32.

CSDF6:  $LSE(CSDF6) = 1.235655+$ . The images used were 12, 17, 29, and 31.

CSDF7:  $LSE(CSDF7) = 1.420816+$ . The images used were 31, 28, 17, 24, 1, and 10.

CSDF7A:  $LSE(CSDF7A) = 6.821821+$ . The images used were 31, 28, 17, 24,

1, and 10. The lesson to be learned from this computation is that given a collection of images, one should do the best job one can in manufacturing an SDF from them.

CSDF8 and CSDF8A were not computed.

These numbers speak for themselves. Suppose one believes, as the past 200 years of science and engineering have empirically verified, that minimizing least squares errors is an intelligent thing to do, and suppose that one accepts the somewhat dubious notion that one cannot manufacture and use the theoretically perfect SDF made from all of the images for some practical reason or other, but instead must make an SDF out of some small subset of them. These computations strongly suggest that a least squares choice and manufacture of an SDF on 5 images always does better than any orthogonalization procedure on 6 images (and usually much better), and that a least squares choice and manufacture of an SDF on 4 images usually does better than most orthogonalization procedures on 6 images (and sometimes much better). They also strongly suggest that the worst orthogonalization procedure is the one which tries to find 6 images which contain the most information about the other images and then take one's SDF to be the theoretically perfect SDF manufactured from these 6 images. I doubt very much if the time consuming computation of SDF8, SDF8A, CSDF8, or CSDF8A would change these empirical conclusions.

#### IV. THE DISCRETIZATION ROUTINES.

Suppose one has computed a theoretical SDF  $h$  in some manner or other, and that one has it in hand in the form of a square matrix ( $512 \times 512$  or  $256 \times 256$  or  $64 \times 64$  or ...) or a vector with many components ( $512^2$  or  $256^2$  or  $64^2$  or ...). One potentially useful way of constructing a physical SDF



begins by discretizing  $h$  so that each pixel assumes one of 256 equally spaced values, adding the magnitude of the largest negative pixel to each pixel ("biasing" the SDF, a very dangerous practice indeed - more on this later), and then displaying and photographing the result in a matrix camera, after proper scaling into grey levels.

One obvious way to discretize the SDF consists merely of taking the range of pixel values, from biggest to smallest, partitioning the range into 255 equal intervals, and then mapping each pixel value to the closest partition point. It is easy to see that this simple practice might give quite bad results. For example, if  $h$  has one pixel  $+1$ , another pixel  $-1$ , and the rest of its pixels nonzero but strictly between  $-1/255$  and  $1/255$ , then  $h_d$ , the discretized version of  $h$ , will have one pixel  $+1$ , another pixel  $-1$ , and the rest will be 0. This  $h_d$  will correlate hardly at all with the  $f_i$ 's. There must be a better way to discretize  $h$ . This simple example suggests that a judicious choice of an upper and lower cutoff and clipping of the pixel values of  $h$  might lead to good results. Two routines were written which attempt to do this in a close to optimal manner. A limited amount of testing has been done so far.

The two programs proceed as follows. Let  $B$  be the biggest and let  $S$  be the smallest pixel value of  $h$ . Let  $CTFU = (3B + S)/4$ ,  $CTFL = (B + 3S)/4$ , and  $MESH = (B - S)/8$ . Successively let  $CTFUT = CTFU + (A \times MESH)$  and  $CTFLT = CTFL + (B \times MESH)$ , where  $A$  and  $B$  assume the values  $-1$ ,  $0$ , and  $+1$  independently. Clip the pixel values of  $h$  above  $CTFUT$  at  $CTFUT$  and below  $CTFLT$  at  $CTFLT$  and discretize the clipped version of  $h$  as follows to obtain  $h_d$ , a discretized approximation to  $h$ : map each pixel value  $Z$  to  $(NINT(255Z/(CTFUT - CTFLT)))((CTFUT - CTFLT)/255)$ . This is not the same

discretization process as indicated in the previous paragraph, for it has the distinct advantage of leaving 0 fixed. This is important! The first program computes and tries to drive to 0  $ERR1 = \|h - h_d\|^2 / \|h\|^2 + LSE(h_d; AVCOR)$ . Here  $AVCOR$  is the average of the 36 numbers  $\langle h_d, f_i \rangle$  and  $LSE(h_d; AVCOR) = |\langle h_d, f_1 \rangle - AVCOR|^2 + |\langle h_d, f_2 \rangle - AVCOR|^2 + \dots + |\langle h_d, f_{36} \rangle - AVCOR|^2$ . In the first pass through the program, 9 values of  $ERR1$  are computed. The new  $CTFU$  and  $CTFL$  are those values of  $CTFUT$  and  $CTFLT$  which simultaneously minimize  $ERR1$ . Then set  $MESH = MESH/2$  and repeat the iteration. Only 8 computations of  $ERR1$  need to be done in the second and subsequent passes through the program. Repeat the iteration until one decides that  $MESH$  is sufficiently small. I stopped this program when  $MESH$  was less than  $(CTFU - CTFL)/512$ .

The second program proceeds in a similar manner, except that it attempts to drive  $LSE(h_d)$  to 0.

Notice that it is a bad idea to try and drive only  $LSE(h_d; AVCOR)$  to 0, for one may only succeed in driving  $AVCOR$  to 0 and perhaps arriving at an  $h_d$  which is the 0 vector. The first summand in  $ERR1$  prevents this from happening, for it demands that  $h$  and  $h_d$  be relatively close. The first term of  $ERR1$  is divided by  $\|h\|^2$  to make both summands have roughly the same order of magnitude.

Each of these programs computes two discretized versions of the SDF. Each of the two versions starts with the same cutoffs, but subsequent scalings may be different. This follows, since in each program the numbers  $LSE(h_d)$  and  $LSE(h_d; AVCOR)$  are computed. But notice that  $LSE(h_d; AVCOR)/AVCOR^2 = LSE((1/AVCOR)h_d)$ .

I had time only to test these routines on CSDF1, CSDF2, and CSDF6. The

following are the results.

CSDF1:  $B = 2.0284 \times 10^{-5}$ ,  $S = -9.1553 \times 10^{-6}$ . The two programs produced 4 discretized versions of CSDF1 which had least squares errors of  $6.4475 \times 10^{-7}$  (CTFU =  $1.9307 \times 10^{-5}$ , CTFL =  $-8.3503 \times 10^{-6}$ ),  $6.3688 \times 10^{-7}$  (CTFU =  $1.9307 \times 10^{-5}$ , CTFL =  $-8.3503 \times 10^{-6}$ ),  $8.4499 \times 10^{-7}$  (CTFU =  $1.8128 \times 10^{-5}$ , CTFL =  $-8.5515 \times 10^{-6}$ ), and  $8.4354 \times 10^{-7}$  (CTFU =  $1.8128 \times 10^{-5}$ , CTFL =  $-8.5515 \times 10^{-6}$ ). If one had done not any clipping, the simplistic approach would have produced two discretized versions of CSDF1 with least squares errors  $4.2690 \times 10^{-6}$  and  $1.6639 \times 10^{-6}$ . So the two routines gave great improvement here. Recall that  $LSE(CSDF1) = 0.0$ .

CSDF2:  $B = 2.1180 \times 10^{-5}$ ,  $S = -7.3345 \times 10^{-6}$ . The two programs produced 4 discretized versions of CSDF2 which had least squares errors of 0.9849 (CTFU =  $6.9367 \times 10^{-6}$ , CTFL =  $-5.5663 \times 10^{-6}$ ), 0.8964 (CTFU =  $6.9367 \times 10^{-6}$ , CTFL =  $-5.5663 \times 10^{-6}$ ), 0.8685 (CTFU =  $1.2324 \times 10^{-5}$ , CTFL =  $-3.0462 \times 10^{-6}$ ), and 0.8898 (CTFU =  $1.2324 \times 10^{-5}$ , CTFL =  $-3.0462 \times 10^{-6}$ ). If one had not done any clipping, the simplistic approach would have produced two discretized versions of CSDF2 with least squares errors 0.8704 and 0.8920. So in this case one of the routines gave a very slight improvement. Recall that  $LSE(CSDF2) = 0.8705$ , so discretization led to something even better. This at first glance seems somewhat contradictory, but discretization is a nonlinear process and produces an SDF which is not a linear combination of the original images.

CSDF6:  $B = 1.8156 \times 10^{-5}$ ,  $S = -5.0980 \times 10^{-6}$ . The two programs produced 4 discretized versions of CSDF6 which had least squares errors of 1.3121 (CTFU =  $6.5047 \times 10^{-6}$ , CTFL =  $-4.8596 \times 10^{-6}$ ), 1.2970 (CTFU =  $6.5047 \times 10^{-6}$ , CTFL =  $-4.8596 \times 10^{-6}$ ), 1.2326 (CTFU =  $1.1911 \times 10^{-5}$ , CTFL =

$-2.4410 \times 10^{-6}$ ), and 1.2762 (CTFU =  $1.1911 \times 10^{-5}$ , CTFL =  $-2.4410 \times 10^{-6}$ ). If one had not done any clipping, the simplistic approach would have produced two discretized versions of CSDF6 with least squares errors 1.2357 and 1.2797. So again one of the routines gave a very slight improvement. Recall that LSE(CSDF6) = 1.2356, again a larger number than given by one of the discretized versions.

These numbers speak for themselves. Though the two routines discussed here need much testing, I believe that they will be of great use in the preparation of more realistic and smaller ( $64 \times 64$  or  $32 \times 32$  or  $20 \times 20$ ) SDF's for actual construction.

#### V. RECOMMENDATIONS, PRACTICES, MALPRACTICES, AND WHAT IS TO BE DONE.

I have already stated my thoughts and opinions on the efficacy of orthogonalization procedures vis-a-vis least squares methods for picking out the best  $p$  from  $m$  images at the end of Section III. I hope that I have demonstrated the usefulness and need for intelligent discretization procedures in Section IV. Let me emphasize once again the care needed in discretization routines to make sure that 0 is mapped to 0, for a little thought will show that grave errors may result if this is not true.

Before turning to more exotic ideas, I first list a number of the more obvious tasks which should be done as a direct extension of the work related in Section III and Section IV.

(1): All of the programs to compute SDF's involves solving systems of linear equations. I did this by writing a straightforward Gaussian elimination scheme from scratch. It seems to have worked very well. However, the process of solving systems of linear equations is known to be numerically very unstable. Minor perturbations in the SDF programs can and

should be made to insure that the solution process is numerically more stable. To be super safe, all of the systems of linear equations should be backsolved to check that the errors made are (hopefully) insignificant.

(2): Notice that all of the routines made worse errors on the clean data than on the tank data. It is obvious that the constant bright background in the dirty data made the correlations much better. However, the errors involving the clean data probably can be (greatly?) reduced if the images are situated vis-a-vis one another in some optimal manner. As a first approximation, the clean images should be translated so that their centroids are at the center of the 256 x 256 arrays. Then all of the routines should be tried again.

(3): SDF's should be made and tested on energy normalized versions of the images. Energy normalization is equivalent to replacing each  $f_i$  by  $f_i / \|f_i\|$ . Most practitioners of the subject insist on this. The usual reason for doing this is to reduce the climatic effects in which the images are located. There are at least two reasons why this practice should be done with a certain amount of caution. Since the SDF will be made from energy normalized images, 32 x 32 pixels in size, it will be looking for energy normalized images, 32 x 32 pixels in size. But an energy normalized 512 x 512 image certainly does not in general have each 32 x 32 subscene proportionally energy normalized. If the object sought is the brightest object in the 512 x 512 scene, then energy normalizing the entire image will leave the object subscene more than proportionally energy normalized, and no harm will result if one is searching for a correlation peak. But if there is a much brighter object, say a fire, in the upper left hand corner of the image, and the object sought is in the lower right hand corner, then

energy normalizing the entire scene perhaps will make the image of the object sought so faint as to be useless. Furthermore, the Schwarz inequality implies that the SDF will have largest inner product with unit vectors that do not look like the objects sought, but instead look like the SDF itself. This difference might be quite pronounced. For the SDF process to work, without further processing, one must make an act of faith that there are no real objects which look more like the SDF than the sought for images themselves.

(4): SDF's should be made and tested on edge enhanced versions of the images. Normally this would involve starting with an image  $f$ , mapping  $f$  to its Fourier transform  $F(f)$ , deleting a suitable disk containing the origin from  $F(f)$ , and mapping the result back to image space by an inverse Fourier transform. All the practitioners of this subject that I have talked to insist on this. The simple example usually given is that if this is not done, then any (uniformly filled in) circle would correlate quite well with a (uniformly filled in) square of roughly the same size. This process also has considerable technical advantages, for if an image  $f$  has already been edge enhanced as above, then  $F(f)$  vanishes at  $(0,0)$  - i.e., the integral of  $f$  over  $R^2$  vanishes. The discretized analogue of this fact is that the sum of the pixel values of  $f$  adds up to exactly 0. Notice that if an SDF  $h$  is made up of a linear combination of such images, it too has the same property. Furthermore, any biasing done to  $h$  then does not change its correlations with any zero mean image. To see this, recall that a biasing of  $h$  involves replacing  $h$  by  $h + h'$ , where  $h'$  is a vector all of whose components have the same constant value, say  $c$ . But if  $g$  is any zero mean image, then  $\langle (h + h'), g \rangle = \langle h, g \rangle + \langle h', g \rangle = \langle h, g \rangle$ , for  $\langle h', g \rangle = 0$  since it

is equal to  $c$  times the sum of the pixel values in  $g$ . Notice that grave errors may be made by blithely biasing SDF's without taking into consideration the mean of  $g$ , for then  $\langle h', g \rangle$  may be quite large. There is a subtle but potentially very serious problem in this circle of ideas. Since the SDF will be made from edge enhanced images,  $32 \times 32$  pixels in size, it will be looking for edge enhanced images,  $32 \times 32$  pixels in size. But, contrary to popular belief, an edge enhanced  $512 \times 512$  image does not have in general each  $32 \times 32$  subscene edge enhanced, in the sense that the sum of the pixel values in this subscene equals 0. For suppose  $f$  is an image such that  $F(f) = 0$  in a disk about the origin in Fourier transform space. Let  $B$  be a square box in image space and let  $I_B$  be its characteristic function; i.e.,  $I_B$  is 1 at points in the box and is 0 at points outside the box. If  $fI_B$  were edge enhanced, then  $F(fI_B)(0,0) = 0$ . But  $F(fI_B)(0,0)$  is the integral of  $f$  over  $B$ , which certainly may be nonzero even if the integral of  $f$  over all of  $R^2$  is 0. For a concrete but somewhat artificial example of this, suppose  $F(f)$  equals  $F(I_B)$  outside of the deleted disk. Then  $F(fI_B)(0,0) = (F(f) * F(I_B))(0,0) = \langle F(f), F(f) \rangle$ , a positive number. This computation does show, however, that the more the original scene is edge enhanced, the smaller is  $F(fI_B)(0,0)$ , but how much edge enhancement is enough to avoid serious errors in the correlation process? These errors might be especially pronounced if one is using a biased SDF.

As I understand it, the dirty tank images had been edge enhanced as above and then biased, so that all their entries were nonnegative, and then discretized into 256 equal parts - hence the bit streams which appeared on the data tape. One way to come close to recapturing the original edge enhanced images would be to take the dirty images, compute the average

pixel value over all pixels between rows 200 to 400 inclusive, and subtract this average pixel value from the same pixels. But then the DeAnza equipment could not have been used to outline the images and toss out the clutter in the background. A somewhat inexact but hopefully fairly reasonable way out of this conundrum is to take each of the clean images, compute the average nonzero pixel value, and subtract this average from each of the nonzero pixels, leaving those pixels with 0 values unaltered. I have already started to do this, having manufactured and tested ZCSDF1, the theoretically perfect SDF on the edge enhanced, clean tank images. All sorts of tests and routines should be run on this most realistic of data up to date.

(5): All possible combinations of (2), (3), and (4) should be done. The most accurate combination probably involves first centering the images, edge enhancing them as best as one can, and then energy normalizing them. Note that simple examples show that edge enhancing and energy normalizing are not commutative operations. For example, suppose an image in its right half is uniformly bright with fuzzy edges and contains a faint object with sharp edges in its left half. First energy normalizing and then edge enhancing would destroy the left hand object and leave an empty scene, while first edge enhancing and then energy normalizing would leave a sharp image of the object on the left and nothing on the right.

(6): Every permutation and combination of SDF's made from the phases of the Fourier transforms of the tank images should be made and tested.

(7): Variants of the discretization routines, in which the process is centered about the average pixel value, not  $(B + S)/2$ , should be run and tested to see if they perform even better than the discretization routines



discussed in Section IV.

(8): While the ideas described to pick the best  $p$  out of  $m$  images to manufacture an SDF work fairly well if  $p$  and  $m$  are not too large, they will not work in a practical sense, if  $m = 100$  and  $p = 10$  say, for then to calculate the analogue of SDF2 would involve finding solutions to (100 choose 10), around  $1.731 \times 10^{13}$ , sets of 10 equations in 10 unknowns, a very large task indeed. Though I am not too optimistic, there might be fairly short computational procedures for finding results close to the theoretically best. Perhaps the random selection and testing of 10 scenes at a time, combined with some sort of gradient technique, would work fairly well.

(9) One can generalize the LSE estimator by replacing each summand by  $w_i | \langle h, f_i \rangle - v_i |^2$ . Think of  $w_i$  as a weight. Usually  $v_i$  will take on the value 0 or 1, but it can be any value one desires. This should have some uses and should be tested in an appropriate setting.

(10) The above suggestions need to be carried out on more realistic data, with  $32 \times 32$  images imbedded into realistic  $512 \times 512$  scenes. In lieu of such data or until it becomes available, the same should be done with simple geometric objects (filled in squares, circles, triangles, quadrilaterals, roughly  $32 \times 32$  in size) embedded into a randomly generated background in a  $512 \times 512$  scene.

Now to more controversial matters. At the present time I know of no scientifically unimpeachable reason why the theoretically perfect SDF should not be used, instead of one made from a small number of pictures, no matter how they are chosen. At first glance it seems implausible that one can do a better job by throwing away information.

One of the alledged reasons for using a smaller number of pictures runs as follows: the more pictures used to make the SDF, the much greater the range of pixel values of the SDF. If true, this would seem to be an important point, for every process to actually manufacture a physical SDF starts directly or indirectly with a discretization process. The more tightly packed the pixel values of the SDF are about their mean, the better should be the discretization process. Limited numerical evidence so far suggests that this is not the case. For example: SDF1,  $B - S = 9.0842 \times 10^{-6}$ ; SDF2,  $B - S = 9.4572 \times 10^{-6}$ ; SDF6,  $B - S = 9.7941 \times 10^{-6}$ . A bit more analysis was done on a CSDF1, CSDF2, and CSDF6. In what follows, NZAV is the sum of the pixel entries divided by the number of nonzero pixels, and NZSTDV is the square root of (the sum of the squares of (each nonzero pixel minus NZAV), divided by the number of nonzero pixels).

CSDF1:  $B - S = 2.9440 \times 10^{-5}$ ,  $NZAV = 1.2104 \times 10^{-6}$ , and  $NZSTDV = 2.7586 \times 10^{-6}$ .

CSDF2:  $B - S = 2.8514 \times 10^{-5}$ ,  $NZAV = 1.4292 \times 10^{-6}$ , and  $NZSTDV = 2.5000 \times 10^{-6}$ .

CSDF6:  $B - S = 2.3254 \times 10^{-5}$ ,  $NZAV = 1.4569 \times 10^{-6}$ , and  $NZSTDV = 2.2537 \times 10^{-6}$ .

Another alledged reason for not using the theoretically perfect SDF, but using instead one made from fewer images, is that the latter SDF will have smaller correlation with clutter in the background. One way in which this might occur is that in the correlation process between the theoretically perfect SDF  $h$  and each of the images  $f_i$ ,  $h$  gives a very high weighting to relatively few pixels in each  $f_i$ , more so than the SDF made from 6, or 5, or 4 images. A statistical analysis of this point can and

should be made. Very careful tests should be done on this question to make sure that the effort needed to generate optimal SDF's made from a minimal number of pictures is justified.

Still another reason for perhaps using SDF's made from a small number of images, rather than the theoretically perfect one, is that their Horner efficiency and their signal to noise ratio may be superior. This can only be checked by manufacture and testing of selected SDF's.

Even if SDF's made from a small number of images have lower correlation with clutter, there still might be several ways to enhance the theoretically perfect SDF. For example, if the theoretically perfect SDF correlates too well on trucks, then remanufacture the theoretically perfect SDF again, this time demanding that it have inner product equal to 1 with each of the tank images, but inner product equal to 0 with a variety of truck images. This is an idea of Hester and Casasent.<sup>11, 12</sup> This idea is already implicit in the edge enhancement practice, for that in effect demands that one manufacture the SDF so that it has inner product 1 with each of the original tank images, but has inner product 0 with a variety of artificial scenes, each of which is the inverse Fourier transform of the Dirac delta function at a well chosen pixel in Fourier transform space. If the theoretically perfect SDF appears to be correlating too well on random, ill defined backgrounds then perhaps one can remanufacture the theoretically perfect SDF, demanding that it be orthogonal to a few randomly generated scenes. Though the probability of this working is low, it is worth trying. I would guess that the price one would pay for these more elaborate schemes is that range of pixel values (the dynamic range) of the SDF would be more diffuse.

Notice that the number of theoretically perfect SDF's is enormous. If  $h$  is a theoretically perfect SDF (i.e., if  $\langle h, f_i \rangle = 1$ ,  $i = 1, \dots, m$ ), then one can add to  $h$  any vector which is in the orthocomplement of the  $f_i$ 's and still obtain another theoretically perfect SDF. Furthermore, they all arise in this manner. So if the  $f_i$ 's are  $d \times d$  images, then the set of theoretically perfect SDF's is a hyperplane in  $R^{d^2}$  of dimension  $d^2 - m$ . This gives one hope that superior SDF's exist.

It is somewhat plausible that the smaller the spread of numbers in an SDF (i.e., the smaller NZSTDV), the better the actual manufactured SDF will perform. This should be true, for the actual manufacture of an SDF hologram, whether optically or on a computer, involves implicitly or explicitly, some sort of discretization. The tighter together the pixel values of the SDF are packed, the better should be this discretization process. I am very, very optimistic that this might be accomplished by judiciously adding to  $h$  a vector  $p$  which is in the orthocomplement of the  $f_i$ 's, with the result that  $\text{NZSTDV}(h + p)$  is close to being a theoretical minimum. This  $p$  will be computed by solving appropriate systems of linear equations. Since it is unlikely that this problem has an actual optimal solution (for the set of potentially useful  $p$ 's certainly is not compact), one will have to settle for a good one. This will require a great deal of computing and numerical experimentation, but it might be worthwhile.

Another potentially useful idea to manufacture superior SDF's is an extrapolation of edge enhancement. Suppose for example, that one manufactures a theoretically perfect SDF on a set of tank images, but finds on experimentation that it unfortunately correlates quite well on truck images. It is plausible that there are certain regions in frequency space

where the Fourier transforms of the truck images and the tank images differ greatly, but that this difference is being overwhelmed by the similarity in the other frequencies. If one actually has these two sets of images in hand, then it is easy to conceive of several computer routines to search for sufficiently large, but not too large, squares, or annuli, or circles in frequency space where this difference is greatest. It is crucial to have a good measure of this difference. As a first suggestion, I would take the following easily computed number. Let  $f_i$  ( $i = 1, \dots, m$ ) be the tank images, and let  $g_j$  ( $j = 1, \dots, n$ ) be the truck images. If  $A$  is a nonempty region in frequency space, let  $I_A$  be its indicator function. That is,  $I_A$  assumes the value 1 at points inside  $A$ , and assumes the value 0 at points outside of  $A$ . One should try and find the  $A$  which minimizes the sum of  $mn$  terms, the typical one which looks like  $\langle F(f_i)I_A, F(g_j)I_A \rangle^2 / \|F(f_i)I_A\|^2 \|F(g_j)I_A\|^2$ . The summand should be dropped, of course, if its denominator is 0.

Finally, and perhaps most importantly, someone with a mathematical mindset should examine the various schemes for making computer generated holograms and see if there is some way in which they can be optimized or improved. Holograms have the very important property that starting with an input image  $f$ , if one manufactures its hologram  $H(f)$ , computes  $F(H(f))$ , and looks at the correct place in the Fourier transform plane, one should see an approximation to  $F(f)$ . This image can never be exact, for it is degraded by all sorts of perturbations. It might be that starting with the correct perturbation  $f'$  of  $f$  and manufacturing  $H(f')$  instead of  $H(f)$ , certain errors might cancel, and that one can find a much better approximation to  $F(f)$  embedded in  $F(H(f'))$  in the appropriate place in the Fourier transform

plane.

#### REFERENCES

1. J. P. Allenbach, Aliasing and quantization in the efficient display of images, Journal of the Optical Society of America, # 6, 69, 1979, pp. 869 - 877.
2. J. P. Allenbach, Representation-related errors in binary digital holograms: a unified analysis, Applied Optics, # 2, 20, 1981, pp. 290 -299.
3. B. Braunecker, R. Heuck, and A. W. Lohmann, Optical character recognition based on nonredundant correlation measurements, Applied Optics, # 16, 18, 1979, pp. 2746 - 2753.
4. D. Casasent, Unified synthetic discriminant function computational formulation, Optical Society of America, # 10, 23, 1984, pp. 1620 - 1627.
5. D. Casasent and V. Sharma, Performance of synthetic discriminant functions for infrared ship classification, SPIE, 422, pp. 193 - 196.
6. D. Casasent, B. V. K. V. Kumar, and V. Sharma, Synthetic discriminant functions for 3-D object recognition, SPIE, 360, 1982, 7 pp.
7. H. J. Caulfield, Role of Horner efficiency in the optimization of spatial filters for optical pattern recognition, Letters to the Editor, Applied Optics, # 24, 21, 1982, pp. 4391 - 4392.
8. H. J. Caulfield and M. H. Weinberg, Computer recognition of 2-D patterns using generalization matched filters, Applied Optics, # 9, 21, 1982, pp. 1699 - 1704.
9. M. H. Hayes, The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform, IEEE Transactions on Acoustics, Speech, and Signal Processing, # 2, ASSP-30, 1982, pp. 140 -154.
10. C. F. Hester and D. Casasent, Multivariant technique for multiclass

- pattern recognition, Applied Optics, # 11, 19, 1980, pp. 1758 - 1761.
11. C. F. Hester and D. Casasent, Intra-class IR pattern recognition using SDFs, SPIE, 292, 1981, pp. 25 - 33.
12. C. F. Hester and D. Casasent, Interclass discrimination using SDFs, SPIE, 302, 1981, pp. 108 - 116.
13. J. L. Horner and P. D. Gianino, Phase-only matched filtering, Applied Optics, # 6, 23, 1984, pp. 812 - 816.
14. K. G. Lieb, M. R. Wohlers, H. K. Hinz, and G. Moyer, Analysis of an optical matched filter guidance system for tactical homing applications, Grumman Research Department Memorandum RM-674J, December 1978.
15. B. V. K. V. Kumar, Errors in optical computation of correlation coefficients, Applied Optics, # 2, 22, 1983, pp. 209 - 211.
16. J. R. Leger and S. H. Lee, Hybrid optical processor for pattern recognition and classification using a generalized set of pattern functions, Applied Optics, # 2, 21, 1982, pp. 274 - 287.
17. A. V. Oppenheim, The importance of phase in signals, Proceedings of the IEEE, # 5, 69, pp. 529 - 541.
18. J. Riggins and S. Butler, Simulation of synthetic discriminant function optical simulation, to appear in Optical Engineering, October 1984.
19. R. H. Squires and J. P. Allenbach, Digital holograms: a guide to reducing quantization and phase encoding errors, SPIE Proceedings, # 4, 437, 1983, pp. 4 - 10.



1984 USAF - SCEE Summer Faculty Research Program

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

STRESS AND AIRCRAFT MAINTENANCE PERFORMANCE

IN A COMBAT ENVIRONMENT

Prepared by: Dr. William D. Kane, Jr.

Academic Rank: Associate Professor

Department and     Department of Management and Marketing  
University:         Western Carolina University

Research Location: Air Force Human Resources Laboratory,  
                         Logistics and Human Factors Division,  
                         Combat Logistics Branch

USAF Research: Robert Johnson

Date: July 20, 1984

Contract No: F49620-82-C-0035

STRESS AND AIRCRAFT MAINTENANCE PERFORMANCE

IN A COMBAT ENVIRONMENT

by

William D. Kane, Jr.

ABSTRACT

This effort examines the impact of stress on the performance of aircraft maintenance in a combat environment. The theoretical literature and research and cases from actual combat situations indicate that performance could be significantly degraded. The literature is not clear on a definition of stress and no practical measure of stress exists. Also, the relationship between stress and performance is not accurately predictable. In addition, maintenance capability will be degraded through psychological casualties. The intensity of the battle will dictate the psychological casualty rate, but most of them will be returnable to duty in three or four days given the proper treatment. The shape of the relationship between combat stress and performance is suggested and recommendations are made as to what additional research might be conducted.

#### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a challenging and constructive summer at the Air Force Human Resources Laboratory, Wright-Patterson AFB, Ohio. He would like to acknowledge the laboratory, particularly the Combat Logistics Branch, for its hospitality and excellent working conditions.

Finally, he would like to thank Mr. Robert Johnson for entrusting him with the preliminary work in this area and for collaboration and guidance. He also acknowledges the many helpful discussions with the branch personnel and particularly Dr. Kristine Yaworsky, a Summer Faculty Research Fellow.

## I. INTRODUCTION

At 0530 hours the message comes into the command post. The wing is being mobilized and sent to its European operational base. Airlift will begin arriving at A+ 48 hours and proceed according to plan. The time has come! International tensions have rapidly escalated in recent weeks, intelligence has reported significant Soviet movements, and the decision has been made to move some of our forces. We should be ready, we have worked and trained hard for the last six months. The maintenance troops have put in a lot of overtime, including weekends, but they are used to that. We have exercised our entire wing mobilization plan and all of the glitches have been worked out. We are ready, initiate the recall!

For the next two days everyone pitches in and works long hours to get everything done. WRSK kits are filled as much as possible and readied for travel. Wheeled equipment is marshalled, test equipment and mockups are packed, and the troops are processed by personnel. The troops have had some time to get their personal affairs settled, besides, they have always known they could leave at a moment's notice. Airlift arrives on schedule, the enroute maintenance team departs, loading proceeds according to plan, the Fighter aircraft are launched, and finally, everyone and everything is loaded somewhere and the last airlift aircraft departs. Quiet settles on the flight line as all of the wings resources are now headed for the bed down location.

At the European base the activity accelerates as stateside resources begin arriving. The first maintenance folks arrive, the first fighter aircraft are due shortly, as are the airlift aircraft with the support equipment, parts, and the rest of the maintenance force. Things are hectic! The incoming fighters are met, serviced and repaired, if necessary. Barracks are assigned, food service is established, parts and equipment are unpacked, test equipment and mockups are set up, communications are established, and late on day four the frag order for munitions upload is received and accomplished. Everything is set.

From a stress perspective you have just delivered a world class fighter to the arena in the worst shape of his career for the fight of his life. The accumulated stress from six months of overtime, four frantic days since recall, sleep deficit, strange and uncomfortable living conditions, disturbed biological rhythms from time zone changes, mess hall food, and the anxiety of expected battle has diminished the total output capabilities of your maintenance force. Worse yet, their stress tolerance threshold for functioning in a hostile combat environment has been lowered. Hegge and Tyner (1982, P.9) indicate that performance can decline 15% just from the effects of the stress of rapid deployment. They also point out that performance reduction stemming from that 15% degradation may result in 100% loss of some tasks. If a team is unloading an aircraft and each team member's performance is off 10% it will take 10% longer to complete the job. However, if the team members are

engaged in a critical, sequential task, such as munitions loading, loss of 15% of any one member's contribution may result in the loss of the entire task as mistakes are made. Dr. (Col.) David Jones, Chief of the Neuropsychiatry Branch at the School of Aerospace Medicine comments that we are all somewhere along the path to combat fatigue. Subject to enough stress we will all succumb and, at least for the short term, be useless or dangerous to the unit's mission. Bringing the maintenance force to combat already well along the stress path steals from that unit's sustainability. If maintenance must launch, service, and repair aircraft in an active combat environment there will be combat fatigue casualties due to the severe stress of that environment, and the higher the stress level troops bring with them to combat the lower will be the unit's sustainability.

Data from World War II and more recent conflicts in the Middle East indicate that in active combat a unit will sustain one psychological casualty for every three physically wounded casualties (Hoey, 1984). In sustained combat, 25% of all casualties may be combat fatigue casualties resulting from an individual's temporary inability to cope with the extreme stress. These data apply to recent conflicts engaged in by combat experienced, well trained soldiers. We can only speculate what the psychological casualty rate might be among Air Force maintenance personnel who have no history, role models, experience, or training with launching, servicing and repairing aircraft in an active combat situation.

Since psychological casualties are generally highest in the early stages of the conflict (Hoey, 1984), and we have brought them to the conflict already fatigued, we could estimate that the percentage of psychological casualties could be quite high. As Dr. Jones comments "In one battle area during Israel's 1973 Yom Kipper War against Egypt, 900 of the first 1,500 Israeli casualties were psychiatric casualties" (Hoey, 1984, P.33). A maintenance unit's sustainability could be badly degraded by stress casualties, and since there is no way to predict the distribution across AFSC's, its capability could be completely lost in some areas.

If stress is such an insidious and debilitating problem, exactly what is it? This discussion will be picked up in Section III, but first a brief discussion of this paper's objectives.

## II. OBJECTIVE

The main objective of this paper is to provide AFHRL/LRC, the Combat Logistics Branch, with an assessment of the feasibility of initiating a work unit to investigate the impact of stress on aircraft maintenance performance in a combat environment. If the project is adjudged feasible I will suggest how to proceed and lay out the rudiments of a research plan. Air Force maintenance units have almost no experience in functioning in a direct, ongoing battle situation where they are carrying out their duties under fire from conventional munitions or threat from chemical/biological agents. Other than isolated instances in World War II, Korea, and Vietnam, Air Force aircraft maintenance has taken place in a relatively

secure rear echelon environment. In almost any NATO scenario it is highly unlikely that aircraft maintenance will take place in a benign environment. As historical evidence indicates that we can anticipate psychological casualties, and probably overall performance degradation, it is imperative that the critical factors of maintenance capability and sustainability receive appropriate attention.

### III. BRIEF REVIEW OF THE LITERATURE

Stress, what is it? Selye (1979, P.12) defines stress as "the non-specific response of the body to any demand." This definition dates back to 1936 and is still reasonably valid, if not very helpful. Non-specific response refers to the fact that the stereotypical stress response can be elicited by any number of stressors, rather than specific stressors. Any stressor, therefore non-specific, produces the stress response. The stress response that Selye outlined and labeled the General Adaptation Syndrom (G.A.S.), emphasizes the evolution of stress in three stages.

1. Alarm reaction. This occurs upon sudden exposure to noxious stimuli to which the organism is not adapted. The reaction has two phases:
  - (a) Shock phase, the initial and immediate reaction to the noxious agent. Various signs of injury such as tachycardia, loss of muscle tone, depressed temperature and blood pressure are characteristic symptoms.
  - (b) Countershock phase, a rebound phase marked by the mobilization of defensive forces. This phase merges into the next defensive phase, during which the adrenal cortex is enlarged and secretion of adrenocorticoid hormones is increased.
2. Stage of resistance. This is marked by full adaptation to the stressor during which symptoms



improve or disappear. There is, however, a concurrent decrease in resistance to other stimuli.  
3. Stage of exhaustion. Since adaptability is finite, exhaustion inexorably follows if the stressor is sufficiently severe and applied for a prolonged period of time. Symptoms reappear and if stress continues unabated, death ensues.  
(Selye, 1979, P.17)

The organism is stimulated, it responds to and adapts to the stressor, and if the stressor is not somehow relieved the organism eventually fails.

Since Selye began his work in 1936 stress has become a very popular topic in a variety of fields. He states that "...we have been able to collect more than 120,000 publications (among them several hundred books) which deal with various aspects of what is now known as the stress concept, not only in virtually all fields of medicine, pathology, biochemistry, and medical jurisprudence, but also in the behavioral sciences and philosophy" (Selye, 1979, P.11). He published an encyclopedia in 1976, Stress in Health and Disease, that contained 7,518 key references. Cohelo and Irving (1981) published a carefully edited, annotated bibliography directed only toward the mental health care and human services field that has 988 entries. The Handbook on Stress and Anxiety (Kutash et al., 1980) has 75 full pages of references. The above are only indicative of the volume of research, yet while much has been learned and written about, the mystery is still not unraveled.

In spite of all of the work, or perhaps because of it, there is currently no agreed upon definition of stress and science is still attempting to fathom its mysteries. Actually, what has been labeled

stress, should more appropriately be labeled arousal, the result of stress. "The identity of the alarm signals that first relay the stress message has yet to be identified." (Selye, 1980, P130). The first mediator is still unknown and what has been observed and measured is the impact of the first mediator. There are almost as many definitions of stress as there are writers about it. In addition to Selye's, among the many are "...stress may be considered as a response to a stressor that induces a change in the individual's ongoing behavioral, physiological or cognitive patterns of functioning; note that no judgment is made as to the valence of the stressful reaction: (Beech et al. 1982, P.10). Another is "Stress will arise whenever the effort mechanism is either seriously overloaded over time or falls altogether short in accomplishing the necessary energetical adjustments." (Sanders, 1983, P.79). Robins et al. (1961) point out a variety of definitions all stemming from the particular approach, discipline, or philosophy employed. However, in order to proceed, this paper requires an operational definition of stress in spite of the fact that it may be imperfect, incomplete or even ultimately false. For the purpose of this paper's focus, the Air Force maintainer under combat conditions, stress is that level of arousal, whatever the stressor, that influences performance.

The issue of performance raises another cloud. The relationship between performance and stress is not well understood and most of the research that has been accomplished is simplistic. (One exception is Ursin et al., 1978.) Much of the work has been done

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC 5.

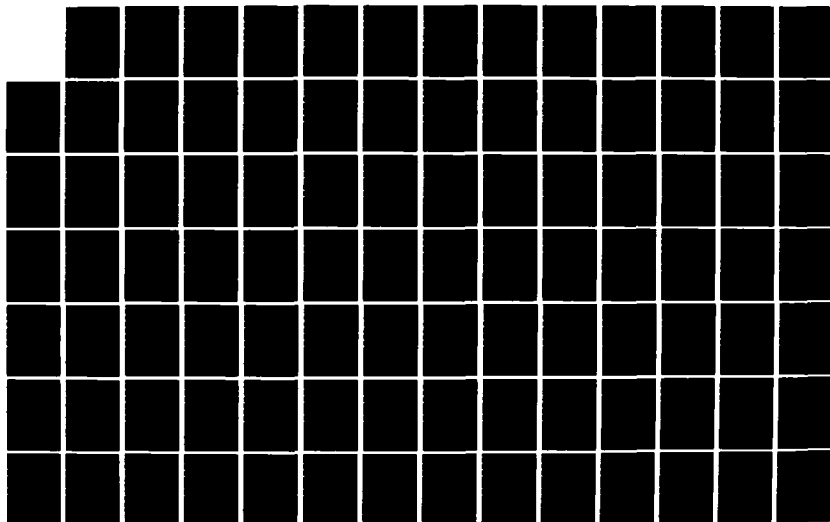
7/13

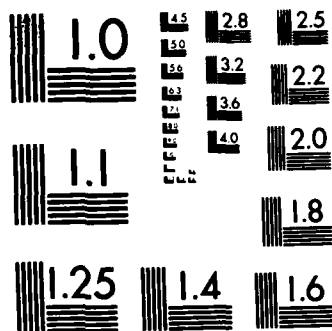
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

with one variable, or very few variables, and most of it in the laboratory. The impact of assumed stressors (e.g. temperature, fatigue, noise) is very complex and any generalization from a stressor to performance is premature. Probably one of the most significant stressors is fatigue, and there is reasonable general knowledge on the limits of fatigue. Yet, the evidence indicates that you cannot predict individual performance because the impact of the fatigue stressor may be moderated positively or negatively by a host of other variables. Stress at the extreme this paper is concerned with, under life and death conditions, is suspected of degrading performance, but to what extent and how soon and for which people is unknown. Maintenance troops have multiple stressors, such as life stress, task demand stress, organizational stress, and combat stress, and at the point of interest, performance on the flight line under attack, prediction to or measurement of individuals is not practical.

Measurement of stress raises other issues. As indicated earlier, what has been measured is arousal and it has been measured in a variety of ways. Unfortunately, the variety of measures produces conflicting results because different stressors cause different reactions and the different reactions measured by different methodologies produce conflicting results. Stress has been measured by self-reports, other paper and pencil instruments, Rorschach tests, urinalysis (catecholamines), blood tests (cortisol, androgens), heart rate, galvanic skin response (GSR), performance, brainwaves, and observation. While there has been good experimental work (e.g.

Ursin et al., 1978 and Bourne, 1969) there is not any useful practical work. Given the variety of stressor sources impacting on the individual there is no way to allocate a portion of the total stress to any stressor or class of stresses as there is no valid or practical way to measure. We can assume that in a combat situation the individual's stress will be very high and that most of his attention will be riveted on that which threatens to destroy him. He may or may not have time (attention) to spend on a maintenance task and if he begins the task he may or may not make errors ranging from trivial to dangerous. But, there currently exists no practical way to accurately measure what is going on or why.

To summarize to this point, we have taken troops into battle who have already "used up" some of their stress tolerance, we have only a tentative operational definition of stress as there is no agreement as to the definition of the concept, an integrated, practical way to measure stress is not available, and it is difficult if not impossible to predict from what we know about stress to performance of maintenance tasks in a combat environment.

The following section introduces a model of stress that represents the process.

#### IV. MODEL OF STRESS

Figure 1 is a model of the stress process. The world in which the Air Force maintainer lives contains many sources of stressors. Stressors impacting on the individual stem from life stress such as family moves, promotion, financial problems, marital issues, role of parent and all of the other factors important to the individual's

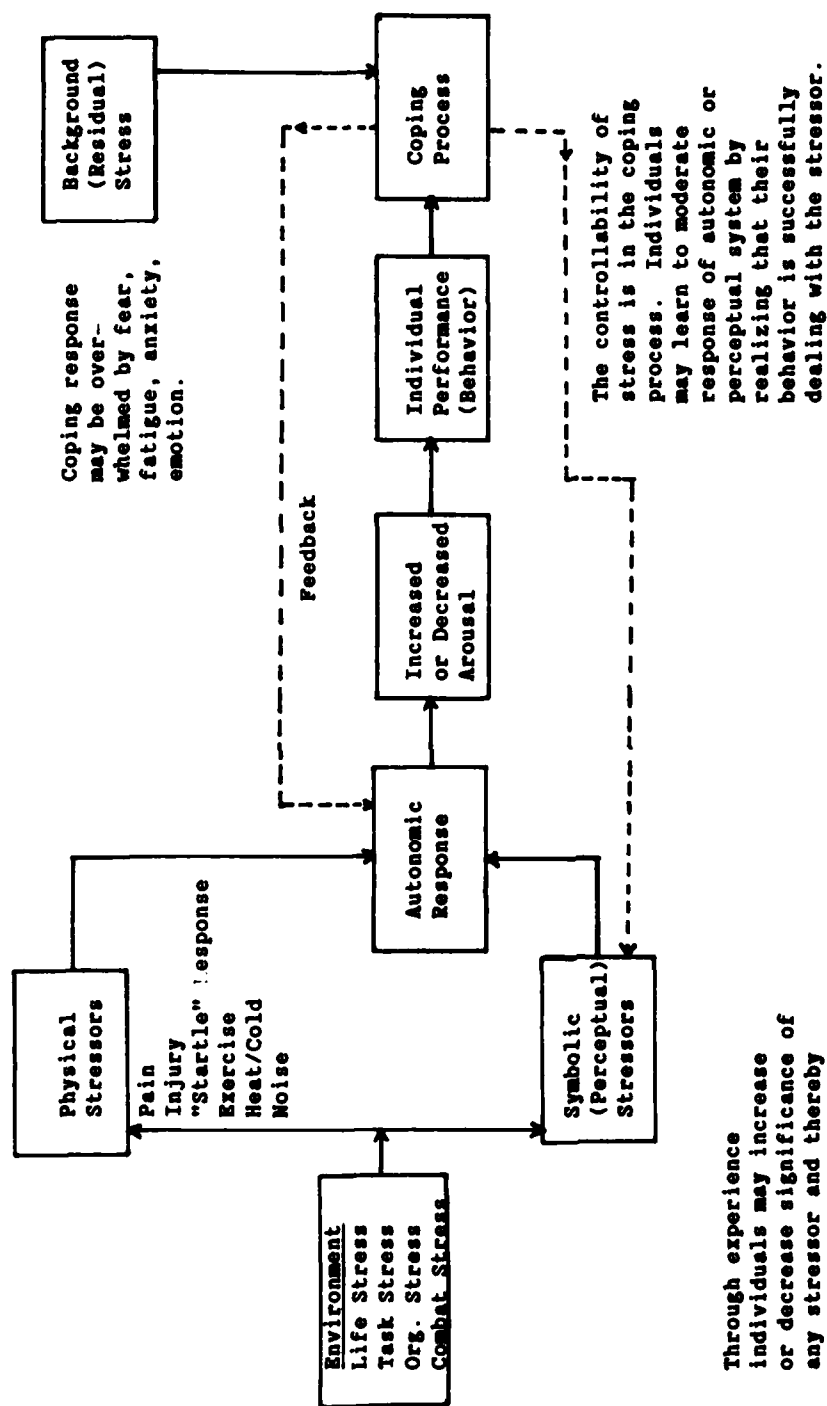


Figure 1  
Model of Stress

living. Task stress derives from the demands of maintenance tasks such as cognitive complexity, effort required, strength required, exposure to heat or cold, variety or lack thereof, and other factors associated specifically with the task. Organizational stress originates in role ambiguity, role conflict, mixed messages, poorly defined goals, perceived sources of inequity, double standards and all those stressors stemming from the organizational context within which the task is accomplished. Combat stress is an extreme case having its source in those events that may kill or maim the maintenance troop in an active combat environment as he strives to service, repair, and launch aircraft. Some of these stressors are relatively pure physical stressors (e.g. when running up steps your heart beats faster, your respiration speeds up, your blood vessels expand, and you begin to perspire) that "turn on" the autonomic nervous system and initiate the G.A.S. The result of "turning on" the autonomic system is some level of increased or decreased arousal (measurable through biochemical urinalysis or blood analysis) which results in some performance (behavior). Science has only recently discovered that some people can consciously control (cope with) this physical autonomic system (e.g. blood pressure, heart rate) by feeding back internally and moderating or shutting down the G.A.S.

Another set of stressors are those whose impact is established by the psychological interpretation of the individual. The stressors are symbolic, they are different than the purely physical stressors



in that the individual "interprets" them as stressful, not stressful, moderately stressful, or somewhere on that continuum. It is within the individual's perceptual process that they are adjudged stressful, and this is why it is difficult to make generic statements about the impact of stress on performance. What individuals interpret as stressful varies from person to person and how each person responds also varies. The autonomic system is "turned on" to some level of arousal, performance (action/behavior) results, and depending upon the strength and appropriateness of coping behaviors, the level of G.A.S. may remain high or the individual can learn to reinterpret the stressor as less threatening because he has engaged successful behaviors and therefore moderates the G.A.S. As Coyne and Lazarus (1980) point out individuals engage in transactions with the multiple environments within which they function whereby they cognitively appraise how they should cope with various stressors. In their model (Coyne and Lazarus, 1980, p. 156) "...coping actively shapes the course of the ongoing person-environment relationship. Rather than a fixed entity that inevitably impinges on the person, much of the environment remains only a potential unit until it is actualized by coping efforts. Environmental influences may shape the constellation of coping efforts that come into play in a stressful transaction, but coping also partially determines which environmental influences are activated and what form they will take." Therefore, while physical stressors are relatively outside individual control symbolic stressors are relatively within individual control and the

variability across people, across tasks, across situations is tremendous. But, if individuals can be taught or self-learn "better" coping strategies, it is possible to moderate symbolic stressors, reduce stress, and theoretically enhance performance.

The model is crude in that it lacks detail, but the amount of detail available would render it useless as a heuristic. It does highlight the process, points out that the process is individualistic, and that to some extent stress management is learnable.

It is appropriate to now put this discussion in the context of the combat environment.

#### V. PERFORMANCE IN COMBAT

"It has been known for a considerable period that soldiers in the heat of battle are unlikely to fire their weapons. After the Battle of Gettysburg in the American Civil War, over 200 of the muzzle-loading rifles were found to have been loaded five or more times without being fired, and one had been loaded 21 times without being fired once." (Idzikowski and Baddely, 1983, p. 128, referencing Walker and Burkhardt, 1965.) Work by Marshall and colleagues in World War II indicated that "...only 15-25 per cent of the soldiers involved in an engagement actually fired their weapons." (Idzikowski and Baddely, 1983, p.126.) Other World War II research indicated that navigational errors increased the closer the bomber got to the target area yet decreased after leaving the target area and heading safely home. Again referencing Walker and Burkhardt (1965), Idzikowski and Baddely (1983) describe research that compared the

ratio of error in combat to error in training. "The results show a decrement of up to 900 per cent as the combat situation becomes more and more dangerous." (p. 129) After reviewing other studies concerned with the relationship between fear and performance Idzikowski and Baddely (1983, pp. 140-141) conclude "Behaviourally, deterioration can be expected in manual dexterity, in sensory motor tasks such as tracking, and in performance of secondary tasks. It is probable that secondary task performance is reduced before central tasks are affected...The evidence suggests that when a situation has induced fear in an individual (as measured by subjective and physiological responses), then a deterioration in the efficiency of performance can be expected, especially in tasks involving sensory-motor skill or divided attention."

If we apply the above conclusions to aircraft maintenance in an active combat environment we can speculate about task performance outcomes. As the intensity of the battle increases individuals' fear/anxiety will increase. As fear/anxiety increases task performance degrades with peripheral tasks being neglected first. We can expect increased mistakes in task accomplishment with decreasing attention paid to secondary tasks such as safety. [It might be possible for safety to become the primary task and the worker forgets about the maintenance task!] At some point all tasks are subordinated to survival. The stress/threat of combat will impact on a maintenance unit's performance.

Cowings (1975) reported on two small Army maintenance support units, one repairing vehicles and one repairing aircraft. Both units experienced unexpected enemy attacks, one including a ground attack, and the attacks had a significant impact on maintenance output (see Figure 2). The initial impact, Phase II, was a decline in performance due to the high state of arousal generated by the shock of the attacks. In Phase III, output increased rapidly past original base line output, Phase I, to an all time high as the troops adapted to the stressful combat environment. In the hyper-efficiency phase, Phase IV, output continues at an abnormal pace as the maintenance troops live off the high stemming from the elevated state of arousal. This period of abnormally high output lasts for some period of time, five to six days in these two cases, and then performance falls abruptly as exhaustion sets in, Phase V. In both of these cases the

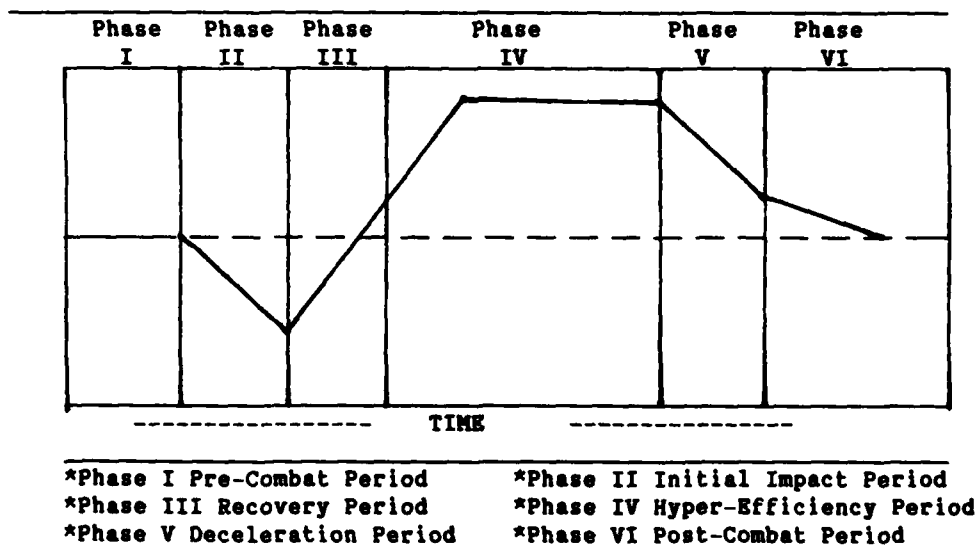


Figure 2  
Impact of Enemy Attack on Maintenance Output  
(Cowings, 1975, p. 90)

commanders immediately recognized the need for rest, battle conditions permitted it, everyone took a day off and had a beer party and output resumed the base line rate, Phase VI. If rest had not been taken it is likely that performance would have continued to decline precipitously. [It must be noted that this composite curve was derived from the subjective evaluations of the commanders of the two units. The amplitude of the curve has only a relative, as opposed to an objective, value, while the duration, time, is more objective but not purely.] It is interesting to note the similarity between this curve and Selye's (1979, p. 16) illustration of the G.A.S. syndrome (see Figure 3). The curves are quite similar and it is not apparent from Cowings' (1975) account that he derived his curve from Selye's. If we accept the curves at face value there is a strong relationship between stress and maintenance output.

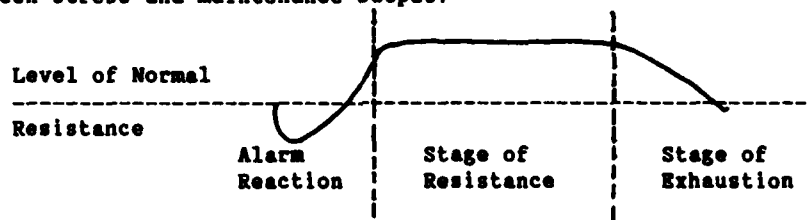


Figure 3  
Resistance During the General Adaption Syndrome  
(Selye, 1979, p. 16)

Another interesting similarity is the shape of Jones' (Hoey, 1984, Jones, 1984) curve of psychological casualties over time, it is the reciprocal of Cowings' (1979) and Selye's (1979). Jones' curve is also based on response to stress and the shape of the response is supportive of the argument that combat stress is going to impact

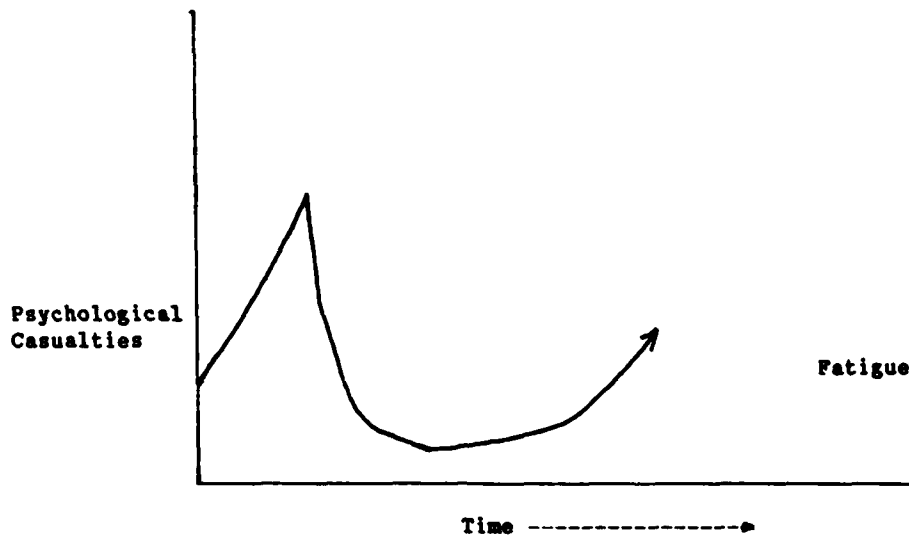


Figure 4  
Psychological Casualty Rate Over Time  
(Hoey, 1984, p. 33)

significantly on maintenance output. Not only is production going to decrease temporarily in the beginning of the conflict we will also have the highest incidence of psychological casualties at the onset. The casualty rate drops off sharply, fairly quickly, but begins to rise again after some time, in the same manner that production peaks for five or six days and then falls abruptly. If some intercession does not occur to moderate or relieve the stress sortie rates will collapse, psychological casualties will rise, and sustainability will be significantly degraded.

Quoting Swank and Marchand (1946), Kern (1966, p. 9) reports their observation of soldiers in the emotional exhaustion phase.

Symptoms which had been developing insidiously now became evident. The soldier was slow witted, he was slow to comprehend simple orders, directions and techniques, and he failed to perform life-saving measures, such as digging in quickly. Memory defects became so extreme that he could not be counted on to relay a verbal order. There was also present a definite lack of concentration on whatever task was at hand, and the man remained preoccupied for the most part with thoughts of home, the absolute hopelessness of the situation and death. This constant dwelling on

death did not indicate a state of fear, but, rather, a certainty that it would occur. The anxious stare, together with tremulousness and generalized hyperactivity was replaced gradually by an emotionless expression, lassitude and listlessness.

While the comments above represent an extreme on the continuum of combat fatigue we might speculate on the quality and quantity of maintenance at other, less extreme points. How many errors will sophisticated weapons systems and munitions tolerate?

Continuing to work from Swank and Marchand's (1946) accounts, Kern (1966) reports their scheme for explaining the development of combat fatigue. The stages are: Phase 1, Initial Combat Adaptation; Phase 2, Period of Maximum Effectiveness; Phase 3, Hyperreactive Phase; and Phase 4, Emotional Exhaustion Phase. Note the similarity to Cowings' (1979) conceptualization (Figure 2) (Cowings does not reference Swank and Marchand) and also note the similarity to Selye's (1979, p. 17) development of the stages of the G.A.S. earlier in these pages. Selye's (1979) three stages are: 1. Alarm reaction; 2. Stage of resistance; and 3. Stage of exhaustion (see Figure 3). While the particulars are not very clear, it is apparent that aside from biochemical measurement there is agreement that the impact of combat stress is going to shape the output of a maintenance unit. What we do not know at the present is what that impact will be, who it will impact, and to what extent it will impact. We do know that it could be severe and therefore devastating.

Another critical factor influencing maintenance output in a battle environment is chemical/biological agents. While no wide-scale use of chemical weapons has occurred since World War I it

appears that the chemical threat may be very high in modern warfare. The Wall Street Journal over the last several months has contained a number of editorials and articles that point out that the Russians train for, are equipped for, and currently use chemical agents. Therefore, maintenance troops, to survive in a chemical environment, will have to function, at least part of the time, in complete head-to-toe protective equipment. Two sources of severe psychological stress are thereby added, the fear and uncertainty associated with a chemical attack and the real problems of performing physical tasks while completely enclosed in protective clothing that you cannot remove or tear or you die.

Cadigan (1982, p. 90) comments that "Casualties will begin to appear, even before real exposure once there is a serious threat perceived of the use of CW by the enemy and some will continue to occur even after the real exposure has passed." Barnard and group (date unknown, p. 123) write that

A powerful source of anxiety is lack of perceptual clarity in the face of apparent danger. Many chemical and all biological weapons are undetectable by the senses, so that there are no warning signs to enable the person to protect himself. Additional uncertainties with biological weapons are the latent periods between infection and illness and the unpredictability of spread through the community. As a result, a person may fear that if he is exposed to these weapons he will not know for certain when he has been infected, how ill he will be, or when the danger has passed. A further confusing factor is that many of the symptoms of illness, especially those involving the gastrointestinal tract, are also symptoms of emotional stress. Thus, if a person develops nausea, vomiting, and diarrhea, he may still not be sure whether he has been infected or not.



History tells us that "gas hysteria" is possible "...in which entire army units would break and run with numerous soldiers manifesting signs of chemical injuries, even though the German Chemical Corps had not employed its arsenal." (Brooks et al., 1983, p. 232). Air Force maintenance troops may have to function in a chemical environment for which they are psychologically unprepared and there are going to be stress casualties as well as the entire range of degraded performance.

Some people will not even be able to tolerate being inside the protective suits and others' task performance will be seriously impaired. Brooks et al. (1983, p. 232) state that "The cocoon-like effect of the protective gear diminishes sensory input and creates a situation analogous to sensory deprivation. Common psychological reactions to sensory deprivation are expected to occur in the chemical battlefield. These potential symptoms include apprehension, paranoia, disorientation, loss of time sense, depersonalization, dissociation, distorted bodily sensations, hallucinations, confusion, and panic." Brooks et al. (1983) conducted a one hour field test with 60 soldiers divided into four groups. There were also 10 moderators/medics. Fourteen of the total 70 participants experienced psychological and behavioral problems while wearing the protective clothing and three of them had to terminate the exercise. Three others had to be taken aside and responded to simple intervention techniques such that they were able to continue with diminished symptoms. These six experienced difficulty in the first ten minutes

of the exercise while the remaining eight had problems in the later part of the exercise. Symptoms observed were anxiety, hyper-ventilation, tremors, confusion, poor judgment, disorientation, difficulty in breathing, panic, visual deficits, fear, obsession, nausea, paranoia, heavy perspiration, memory loss, and claustrophobia. Not all 14 soldiers who had problems had all of the problems. The point is that 20 per cent of the personnel experienced difficulty just being in the the suit knowing that it was only for one hour and there were no chemical agents present except a light smoke to indicate attack. Generalizing these results to an Air Force flight line under chemical attack raises real concern about levels of performance and how they will be impacted by fear of chemical attack as well as performing in protective clothing. Symptoms such as confusion, panic, poor judgment, visual deficits, memory loss, and disorientation do not bode well for the performance of critical, cognitively complex, or even sequential tasks.

To summarize this section and tie it back to the previous section, stress will impact performance under combat conditions. It is not clear exactly what stress is, there is no practical way to measure it, it is not possible to accurately predict from stress to performance, performance may be severely impacted by combat conditions, and the whole thing is confounded by the threat of maintaining weapons systems in a lethal chemical environment. The bottom line is that it is possible that stress will seriously limit the capability of a maintenance unit to produce the expected number of required sorties.

## **VI FEASIBILITY**

Should the Combat Logistics Branch (AFHRL/LRC) undertake research into the impact of stress on maintenance performance in a combat environment? Yes. The issue is of central importance to readiness and sustainability and in my opinion the research philosophy, experience, and expertise of the branch make it uniquely qualified to investigate at least parts of the problem. There are some areas of the research that need to be done by others, or cosponsored (e.g., medical, biochemical), but there is much that LRC can do. The following section provides specific recommendations.

## **VII RECOMMENDATIONS**

There are five recommendations, listed in the order in which they should be undertaken. The criterion for the sequence of the list is the time required before something useful can be produced that will enhance maintenance performance in combat. The five recommendations, which will be discussed individually, are:

1. Develop two education programs for stress management in the Air Force maintenance environment with emphasis on reducing the impact of stress on combat maintenance. One program to focus on all supervisory personnel and one program to begin to condition maintenance personnel for functioning in a combat environment.
2. Conduct a study of Air Force maintenance personnel, sortie producing AFS's, comparing them against all other Air Force personnel to determine if there is a higher incidence of

stress related disease among maintenance types. This would be designed to determine in a macro fashion if maintenance is a more stressful environment.

3. Devise an Index of Organizational Stress. It might be possible to develop a series of indicators that would make up an index of organizational stress that would provide a gross measure of a wing's cumulative stress.
4. Investigate the possibility of teaching individuals better coping strategies whereby they conduct more competent transactions with their environments.
5. Begin experimental work on the flight line to develop multiple measures of stress that may eventually result in a practical measure of stress. Also investigate the relationship of stress to performance in the complex context in which maintenance takes place.

The first recommendation is designed to provide Air Force maintenance personnel with new information about the combat context, and about themselves, that will enable them to modify symbolic stressors through the coping process (see Figure 1). At the same time it is necessary to teach maintenance managers to better recognize and manage stress to preclude unnecessary performance degradation. Very few maintenance managers have experience in fixing airplanes in an active combat environment and their handling or mishandling of the early signs of stress can make the difference between a productive maintenance man and a psychological casualty. Their skill in stress

management in the peacetime environment can not only enhance productivity, it should also reduce the level of cumulative stress troops take with them into combat and thereby increase their ability to function in an extremely stressful environment.

Exactly what these two programs should look like is not known at this time. What can be suggested is what they should not look like. They should not be a typical lecture method presentation that personnel are exposed to once in their Air Force careers or annually at a required commanders call for fifteen minutes. Treated stereotypically, combat stress training would quickly be lost in the background noise of on-going events. More than likely if it were presented in the usual manner Air Force maintainers would perceive the program as just another one of those distractors (stressors) that take time away from maintenance and they would be correct! The type of exposure needed to incorporate stress awareness and management into supervisory skills and personal stress management requires far more than passive lecturing. Information transfer, pure intellectual awareness, will not bring into practice the necessary sensitivity and techniques for successful stress management. The typical indoctrination program may actually create more harm than good (Kern, 1966). Kern (1966) argues that all of our experiences, training, and living create and reinforce two attitudes, one called confidence and one called despair. Successful and "good" things reinforce confidence while failure and "bad" things reinforce despair. He argues that the typical orientation tells you

all of the bad things that can happen, often with graphic pictures (e.g., pictures of the effect of venereal disease and frost bite), but seldom provides confidence building strategies for avoiding despair oriented outcomes. Typical orientation programs reinforce despair rather than confidence and therefore defeat the purpose for which they are intended. If combat stress management is presented as a traditional orientation program the impact would likely be negative as the recipient finds out that he could get killed in combat. What must be emphasized is that the Air Force maintainer can function productively in a combat environment and here are some ways that, with practice and use, will help you deal with fear and other stressors.

The "off-the-shelf" packages carried in stock by many consultants and trainers will not work either. They are aimed at a different audience, in a different context, and most of them are not demonstrably effective. They are superficial, quick, and typically designed to meet the needs of the consultant/trainer rather than the client. The questions of what should be delivered (curricular content), when, where, how, and by whom are that, questions. Premature or self-serving answers will only insure that money is spent on something that has a high probability of not working. It may be that either or both of these programs delivered "properly", that result in high impact, will be expensive. But cheaper, diluted versions would be useless and therefore wasteful. There is probably some threshold above which the training is effective and produces change in

behavior. Below that threshold the training is useless as it has little or no impact and behavior remains as it was before.

The objective should be to provide training that enhances the stress management capability of supervisors and individuals such that they function more productively in both a peacetime and combat environment. Precisely how that objective should be met for Air Force maintainers is, at this time, a research question. It may also be necessary to tie this in with recommendation number four.

The second recommendation, to compare the incidence of stress related disease between the maintenance population and the rest of the Air Force, is a macro, as opposed to a micro, approach to the problem. Ideally, the investigator would identify sortie producing AFS's and compare the incidence of stress related disease (e.g. hypertension, heart disease, gastro-intestinal ailments) to all other AFS's. Unfortunately, such data probably do not yet exist. Base level hospitals report disease by category for admitted patients, but the data are not easily traceable to AFS. Neither does base level data report sick call incidences. I also suspect that a lot of self-treatment (e.g. Malox for gastro-intestinal) takes place and that data would be difficult to obtain. Testing the hypothesis that the maintenance environment is more stressful than other Air Force environments could require data that does not yet exist.

Several possibilities should be pursued, however. As stress related disease may not manifest itself until retirement or after, the retired rolls might contain useful data. Line 68 of the retired

airmen file (line 75 for officers) lists percent disability at retirement. Lines 70, 71, 72 (airmen) and lines 77, 78, 79 (officer) are VA diagnostic codes that probably describe the disability and it might be possible to assess incidence of stress related disease. At this time I do not know if the retired file is updated with medical data subsequent to retirement.

The most likely avenue is one discovered in a telephone conversation with Col. Detreville (AV 240-2002) of the Air Force Occupational and Environmental Health Laboratory at Brooks AFB, TX. The Lab is currently developing a Standardized Occupational Health Program (SOHP) and a Comprehensive Occupational Health Program (COHP) and they have talked about including stress data at some point. However, at the present the programs do not include stress data. Col. Detreville also mentioned an Air Force Heart Program sponsored by the Air Force Surgeon General's Office that is looking at, among others, stress factors.

These avenues, and others yet unidentified, should be pursued as sources of data to test the stressful environment hypothesis. It may be necessary to get the answer to this question to provide the support to carry out recommendation number one.

The third recommendation, to devise an index of organizational stress, is another macro approach. It should be possible to develop a set of indicators that cumulatively yield an organizational stress index. Of particular interest would be an index sensitive enough to reveal sudden changes in overall unit stress. What the indicators



are is an empirical question, but some of the possibilities are: AWOL rate, sick call rate, incidence of disciplinary action, quality control failure rate, requests for transfer, spouse abuse, child abuse, alcohol and other drug abuse, sale level of self-treating antacid in BX, racial conflict increase, rise in insubordination, incidences of maintenance malpractice, confirmed diagnoses of stress related diseases, flight surgeon's assessment of unit's mental health, extent of overtime, average time since last move, divorce rate, perception of leadership effectiveness at numerous levels, accident rate (on and off duty) and others not listed here. The point of the index is to provide an indicator of a unit's cumulative stress level as a warning that intervention techniques are needed to stop the degradation of the unit's readiness and sustainability. While not precise, I believe it possible to construct such an index that could provide an early warning to dysfunctional stress increases.

The fourth recommendation, teaching individuals better coping skills, would be an ambitious project. It might also need to be tied to recommendation number one. Current stress management programs are aimed at teaching professionals, generally managers, better coping skills. The impact of this training in terms of substantive changes in behavior is questionable. The theory and the practical techniques are there and seem to work, but most people lack the commitment to change their behavior and incorporate stress management practices into their way of life. Therefore, they are intellectually exposed

to the material, agree that they are stressed, would like to be less stressed, but do not make the behavioral commitment to change. The training happens but nothing changes.

Work is just beginning on stress in the blue collar work force, but unless different presentation techniques or content or whatever are utilized the results will be the same, minimal. However, referring back to Figure 1, the greatest possibility for moderating stress is through the coping process whereby symbolic stressors can be modified. Coyne and Lazarus (1980) have a transactional theory of stress whereby individuals engage in transactions with their environment. The quality of these transactions can be improved through better coping skills. Individuals can reduce their experienced stress by becoming more proficient in environmental transactions (reinforce confidence - Kern (1966)). How to go about, on a large scale, teaching maintenance personnel improved coping behaviors that have a measureable impact on performance is an interesting question, one that I believe it is imperative to begin work on, but one for which I do not have an answer.

The fifth, and last, recommendation, work on experimental measurement of stress, has the least probability of providing a "useful" outcome. The results could add to scientific knowledge about stress but still not help with the practical problem of a simple stress measurement. The effort would have to be multidisciplinary and use multiple measures. Different stressors cause different reactions and the interaction of multiple stressors causes yet different results.

Results across individuals vary greatly because of individual differences in symbolic stressors and, perhaps, stress resistance. This effort would be experimental, long term, complex, and problematic. The literature indicates that you can make a career out of trying to measure stress. In almost 30 years of research science has learned a lot about various aspects of stress, but there is still much to be learned. As a scientific effort it would be worthwhile. As something that would have a practical payoff, it is not likely.

#### VIII RECOMMENDED RESEARCHERS

The following individuals are suggested as initial contacts for the recommendation indicated. I have not talked with any of these people so I do not know of their interest. They are recommended on the basis of published literature.

Recommendation one. Dr. David Jones said that as far as he knew no one was working on stress management in an aircraft maintenance environment. As author of this paper, Dr. William Kane will submit a follow on proposal to AFOSR to follow up on recommendation number one.

Recommendation two. I suggest using someone familiar with the aircraft maintenance work and have him or her work with the occupational health and heart disease medical people to find or devise the appropriate data. The statistical design should be relatively straightforward once the data are available.

Recommendation three. Dr. I.G. Sarason, Department of Psychology, University of Washington, Seattle, Washington, Phone (206) 543-2640 was working on a Navy sponsored project titled "Personnel Technology: Effects of Life Stress and Coping Skills on Performance and Organizational Effectiveness" in May, 1981. From the abstract I read it appears that his work might be applicable here. Dr. Sarason is a big name in stress research.

Recommendation four. Dr. Richard S. Lazarus of the University of California at Berkeley should be the person contacted here. He has done a lot of work over a long period of time and is an established figure in the field. I am not sure that he has actually tried implementing his theory in a practical way, but he has the most appropriate theory.

Recommendation five. If the effort is attempted it will need a multidisciplinary team. The team should consist of a research M.D., a psychologist, a statistician and research designer who could also be the computer specialist, a biochemist (at least as a consultant), someone familiar with maintenance, and a saint to get them all to work together.

NOTE: The folks at HumRRO, The George Washington University should be contacted about all five areas. They have a long history of Army combat research, e.g., Task Fighter series.

#### IX CONCLUSION

Are Air Force maintenance people going to be stressed to the breaking point in a combat environment? Yes, some of them are. But, if

handled properly, most of them can be returned to duty in three or four days. The Israeli's demonstrated that approximately 75% of their psychological casualties could be returned to combat with proper water, rest, food, and treatment. Also, the greater the intensity of the battle the higher the casualty rate and the higher the incidence of psychological casualties. For every three physically wounded there will be one psychological casualty. They will not be cowards, they will be victims whose coping mechanisms have temporarily broken down, they are overwhelmed. There will also likely be more psychological casualties at the onset of combat than later on. In terms of maintenance capability it will be reduced, at least temporarily, by psychological casualties.

Is the stress of combat going to impact on sortie generation rates? Probably, but the exact impact is not known at this time. Limited data indicate that combat max efforts will only last about five days. Dr. David Jones concurs with this assessment, but comments that it is more of a gut feel than objective data. After that time period exhaustion sets in and productivity declines sharply. Combat stress will impact maintenance performance incrementally as well as catastrophically. Mistakes will be made, tasks will take longer, some tasks may be ignored or forgotten as individuals concentrate on personal safety. Functioning in a full chemical suit will compound matters as the suits in and of themselves are significant stressors. The exact impact of combat stress on maintenance performance is unknown, but it is likely to be significant.

Can anything be done to moderate the impact of combat stress on maintenance productivity? Perhaps. Teaching Air Force supervisors to identify stress symptoms and then engaging in appropriate intervention techniques could reduce the negative impact of stress. Teaching individuals better coping skills could provide them with additional psychological strength to resist performance degradation due to combat stress. Managing stress levels in peace time so as not to bring troops into combat already well along the stress-performance curve could enhance sustainability. The matter of what can be done to moderate the impact of combat stress on maintenance performance is a research question and one that needs immediate attention.

It is difficult to do justice to a question of this scope and importance in a short, ten week period. I have sampled a variety of literatures, attempted to integrate them as their central themes pertain to the question at hand, and then projected that integrated perception into several suggestions as to how research into the question might proceed. It is my hope that, while I have not provided answers, I have added a soupcon of clarity to the questions.

#### SELECTED REFERENCES

Alluisi, Earl A., "Stress and Stressors, Commonplace and Otherwise," in Earl A. Alluisi and Edwin A. Fleishman (Eds.) Human Performance and Productivity: Stress and Performance Effectiveness, Hillsdale, N.J., Lawrence Erlbaum Associates, 1982.

Alluisi, Earl A., and Edwin A. Fleishman (Eds.), Human Performance and Productivity: Stress and Performance Effectiveness, Hillsdale, N.J., Lawrence Erlbaum Associates, 1982.

Austin, F.H., Jr., "A Review of Stress and Fatigue Monitoring of Naval Aviators During Aircraft Carrier Combat Operations: Blood and Urine Biochemical Studies," in P.G. Bourne (Ed.) The Psychology and Physiology of Stress with Reference to Special Studies of the Vietnam War, New York, Academic Press, 1969, pp. 197-218.

Beech, H.R., L.E. Burns, and B.F. Sheffield, A Behavioral Approach to the Management of Stress, New York, John Wiley & Sons, 1982.

Berkum, M.M., "Performance Decrement Under Psychological Stress," Human Factors, 6, 1964, 21-30.

Bernard, Viola W. (Coordinator) and group, "Psychosocial Consequences of Chemical and Biological Weapons" prepared at WHO's request, Division of Community and Social Psychiatry, Columbia University, New York, Annex 6, pp. 121-127, date unknown.

Bhagat, Rabi S., "Effects of Stressful Life Events on Individual Performance Effectiveness and Work Adjustment Processes Within Organizational Settings: A Research Model," Academy of Management Review, vol. 8, no. 4, Oct 1983, pp.660-671.

Bourne, Peter G. (Ed.), The Psychology and Physiology of Stress with Reference to Special Studies of the Vietnam War, New York, Academic Press, 1969.

Brooks, Franklin R., Donald G. Ebner, Stephen N. Xenakis, and Paul M. Balson, "Psychological Reactions During Chemical Warfare Training," Military Medicine, vol. 148, March 1983, pp. 232-235.

Cadigan, Francis C., Jr., Col., M.D., "Battleshock - The Chemical Dimension", Journal of the Royal Army Medical Corps, 1982, 128, 89-92.

Clemes, Harris, "Relationship of Military Environments to Stress and Performance," U.S. Army Medical Research and Development Command Contract # DA-49-193-MD-2637, November, 1971 (AD 733933).

Coelho, George V. and Richard I. Irving, Coping and Adaptation: an annotated bibliography and study guide, National Institute of Mental Health, 1981.

Cowings, John S., "Reaction of Combat Service Support Troops Under Stress: The Small Maintenance Support Unit in a Combat Environment," U.S. Army Command and General Staff College Thesis, June 1975 (AD B006734).

Coyne, James C. and Richard S. Lazarus, "Cognitive Style, Stress Perception, and Coping," in Irwin L. Kutash, Louis B. Schlesinger, and Associates (Eds.), Handbook on Stress and Anxiety, San Francisco, Jossey-Bass, 1980.

Davis, S.W., "Stress in Combat," Scientific American, 1956, 194(3), 31-35.

Drabek, Thomas E., "Theory and Methods in the Study of Organizational Stress," Final Scientific Report, U.S. Air Force Office of Scientific Research, Nov, 1968 (AD 678929).

Eden, Dov, "Critical Job Events, Acute Stress and Strain: A Multiple Interrupted Time Series," Organizational Behavior and Human Performance, vol. 30, no. 3, Dec 1982, pp 312-329.

Ellertsen, Bjorn, Tom Backer Johnsen, and Holger Ursin, "Relationship between the Hormonal Responses to Activation and Coping" in Holger Ursin, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Hamilton, Vernon and David M. Warburton (Eds.), Human Stress and Cognition, New York, John Wiley & Sons, 1979.

Hansen, Jan R., Karl F. Stoa, Arnoldus Schytte Blix, and Holger Ursin, "Urinary Levels of Epinephrine and Norepinephrine in Parachutist Trainees," in Holger Ursin, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Hegge, Frederick W. and C. Frederick Tyner, "Deployment Threats to Rapid Deployment Forces," Walter Reed Army Institute of Research, Washington, D.C., December 1982 (WRAIR NP-82-2).

Hockey, Robert (Ed.), Stress and Fatigue in Human Performance, New York, John Wiley & Sons, 1983.

Hoey, Brian, "Wars of the Mind," AIRMAN, May 1984, pp. 33-36.

Hogan, Robert and Joyce C. Hogan, "Subjective Correlates of Stress and Human Performance" in Earl A. Alluisi and Edwin A. Fleishman (Eds.), Human Performance and Productivity: Stress and Performance Effectiveness, Hillsdale, N.J., Lawrence Erlbaum Associates, 1982.



Idzikowski, Chris and Alan D. Baddeley, "Fear and Dangerous Environments" in Robert Hockey (Ed.), Stress and Fatigue in Human Performance, New York, John Wiley & Sons, 1983.

Jamal, Muhammad, "Job Stress and Job Performance Controversy: An Empirical Assessment," Organizational Behavior and Human Performance, vol. 33, no. 1, Feb 1984, pp. 1-21.

Jones, David Dr. (Col.), Chief of Neuropsychiatry Branch, Aerospace School of Medicine, Brooks Air Force Base, TX, Personal interview June 12-13, 1984.

Kern, Richard P., "A Conceptual Model of Behavior Under Stress, With Implications for Combat Training" prepared for Office, Chief of Research and Development Department of the Army by the George Washington University Human Resources Research Office (HumRRO), June 1966 (AD 637312).

Krahenbuhl, Gary S. and Joseph Harris, "Biochemical Measurement of the Human Stress Response," for the Air Force Human Resources Laboratory, by Human Performance Laboratory, Arizona State University, Tempe, Arizona, March 1984 (AFHRL-TR-83-40).

Kutash, Irwin L. and Louis B. Schlesinger and Associates (Eds.), Handbook on Stress and Anxiety, San Francisco, Jossey-Bass, 1980.

Lester, James T., "An Annotated Bibliography of DoD-Funded Reports Concerning Psychological Stress: 1950-1978," Office of Naval Research, September 1979 (AD A080348).

Levine, Seymour, Joanne Weinberg, and Holger Ursin, "Definition of the Coping Process and Statement of the Problem," in Holger Ursin, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Levine, Seymour, "Cortisol Changes Following Repeated Experiences with Parachute Training" in Holger Ursin, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Malkoff, Donald B., "Biotechnology Predictors of Physical Security Personnel Performance: Cerebral Potential Measures Related to Stress," Navy Personnel Research and Development Center, San Diego, CA, February 1984 (TR 6704).

Mountcastle, Vernon B. (Ed.), Medical Physiology, St. Louis, the C.V. Mosby Company, 13th ed., 1974.

Nugent, William A., "Biotechnology Predictors of Physical Security Personnel Performance: II. Survey of Experimental Procedures to Assess Performance Under Stress," Navy Personnel Research and Development Center, November, 1983 (NPRDC Special Report 84-9).

Poulton, E.C., Environment and Human Efficiency, Springfield, Illinois, Charles C. Thomas, 1970.

Robins, James E., James M. McKendry, and Paul M. Hurst, "Task-Induced Stress: A Literature Survey," HRB-Singer, Inc., October 27, 1961 (AD 667272).

Robinson, Eleanor R.N., "Biotechnology Predictors of Physical Security Personnel Performance: I. A Review of the Stress Literature Related to Performance," Navy Personnel Research and Development Center, June 1983 (NPRDC TN 83-9).

Sanders, A.F., "Towards a Model of Stress and Human Performance," Acta Psychologica, 1983, 53, 61-97.

Sarason, Irwin G., and Charles D. Spielberger (Eds.), Stress and Anxiety, Washington, D.C., Hemisphere Publishing Corporation, 1979.

Selye, Hans, "The Stress Concept and Some of its Implications," in Vernon Hamilton and David M. Warburton (Eds.), Human Stress and Cognition, New York, John Wiley and Sons, 1979, pp. 11-32.

Selye, Hans, "The Stress Concept Today" in Irwin L. Kutash, Louis B. Schlesinger and Associates (Eds.), Handbook on Stress and Anxiety, San Francisco, Jossey-Bass, 1980, pp. 127-143.

Ursin, Holger, "Activation, Coping, and Psychosomatics," in Holger Ursin, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Ursin, Holger, Eivind Baade, and Seymour Levine (Eds.), Psychobiology of Stress, New York, Academic Press, 1978.

Vossel, Gerhard and Werner D. Froehlich, "Life Stress, Job Tension, and Subjective Reports of Task Performance Effectiveness" in Irwin G. Sarason and Charles D. Spielberger (Eds.), Stress and Anxiety, Washington, D.C., Hemisphere Publishing Corporation, 1979, pp. 199-211.

Wallen, Vincent, "Background Characteristics, Attitudes, and Self-Concepts of Air Force Psychiatric Casualties from Southeast Asia," in P.G. Bourne (Ed.), The Psychology and Physiology of Stress with Reference to Special Studies of the Vietnam War, New York, Academic Press, 1969, pp. 167-196.

Wilkins, Walter L., "Psychophysiological Correlates of Stress and Human Performance," in Earl A. Alluisi and Edwin A. Fleishman (Eds.), Human Performance and Productivity: Stress and Performance Effectiveness, Hillsdale, N.J., Lawrence Erlbaum Associates, 1982.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

FINITE ELEMENT ANALYSIS OF CENTRIFUGED CONCRETE CULVERTS

Prepared by:	Dr. Yong S. Kim
Academic Rank:	Assistant Professor
Department and University:	Department of Civil and Mechanical Engineering McNeese State University
Research Location:	Air Force Engineering and Services Center, Engineering Research Division, Airbase Survivability Branch
USAF Research:	Dr. Paul Y. Thompson
Date:	September 4, 1984
Contract No.:	F49620-82-C-0035

# FINITE ELEMENT ANALYSIS OF CENTRIFUGED CONCRETE CULVERTS

by

Yong S. Kim

## ABSTRACT

This study shows that finite element analysis can be a useful tool in examining the validity of the results of centrifugal model testing as long as the constitutive relationships for soil are representative of real ones. The quasi-theoretical results obtained from the finite element computer code, CANDE, are compared to the measured behavior of a 4.3-inch diameter instrumented culvert subjected to a gravity force of 35g in the centrifuge. A comparison of the quasi-theoretical and actual behavior shows good correlation. Two soil models, a linear model and Duncan's hyperbolic model, were used to show the influence of the constitutive relationships for soil on the culvert response in the finite element analysis.

#### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Engineering and Services Center, Tyndall AFB, Florida. He would like to acknowledge the Center, in particular the Airbase Survivability Branch, for its hospitality and excellent working conditions.

Also, he would like to thank Dr. Paul Y. Thompson for suggesting this area of research and for his collaboration and guidance. Finally, he would like to acknowledge many helpful discussions with Capt. Paul L. Rosengren.

## I. PURPOSE:

Soil-structure interaction problems have long been areas of interest and concern for geotechnical engineers. Although rapid progress has been made in solving these problems, their application to concrete culverts embedded in embankments has been very slow. Due to an increased number of concrete culvert projects in industry, state highway departments, and federal agencies in the United States during recent years, research could improve the design/analysis of concrete culverts and result in better structural stability, and significant time and cost savings.

## II. OBJECTIVES:

One currently used design method for buried culverts, originally developed by Marston (9) early in the century, has been modified by Spangler (10) and Costes (1). According to this method, the design of buried culverts in a soil fill is dependent on two empirical factors, settlement ratio and load factor. However, primary shortcomings of this method are: (1) the settlement ratio is not known before the culvert is built; and (2) the designer must predict this parameter from a few past experimental observations or by engineering judgment. Therefore, there is a need for more rational design methods on a sounder theoretical basis.

Perhaps the most ideal approach for obtaining information for the development of new design methods is full-scale model testing. A full scale model with the necessary instrumentation (i.e. soil stress meters, settlement gauges, and strain gauges, etc.) could give the best results

for estimating prototype behavior of buried culverts.

Unfortunately, full-scale model testing has serious major drawbacks: mainly, cost and time of construction and operation. Because of these reasons, the centrifugal modeling technique is becoming a favorite testing method in geotechnical engineering. This technique offers not only a comparatively inexpensive and easy way of obtaining essentially the same data as could be obtained from field tests. It also provides opportunities for studying the effect of individual parameters, and collecting more data than field tests usually yield.

However, at present, there is limited capability for soil to be added to, or removed from, a model in centrifugal flight. Thus a typical test of a model embankment involves subjecting that embankment to a gradually increasing centrifugal field, while maintaining it at its final shape. Although the models can be made large enough to include blocks of undisturbed and compacted soil which are representative of field conditions (Rowe [11]), an objection could be raised that the stress paths traversed by soil elements during the test are unrepresentative of those which apply in the field.

Therefore, this necessitates computational methods which examine the validity of such objections. Using finite element techniques, mathematical models can be created which model both the laboratory and the field situations, and allows the two to be compared. As long as constitutive equations for the soil used in the mathematical model are representative of real ones, valid comparisons should be possible.

Therefore, the objectives of the study are:

- (1) to investigate the influence of the constitutive equations

for soil on the culvert response in finite element analysis, and

- (2) to compare the results of the finite element analysis to the results of the centrifugal model testing of the concrete culvert.

### III. CENTRIFUGAL MODEL STUDY:

The centrifugal model testing of the buried concrete culvert in sand was performed by James and Larsen (3) at Cambridge University, and only brief descriptions for the model testing procedures are summarized here.

A strong box, 17.8 X 15.7 X 5.9 inch rectangular box was mounted on the centrifuge. The box consisted of a frame made of hollow sections welded together, the frame was bolted to a back plate of mild steel and the front of the box was a slab of Perspex, which permitted observation of the model.

After mounting the strong box, the pipe, which was cast of micro-concrete (i.e. concrete with scaled down aggregates), was placed 1.2 inches away from the bottom of the box. The load cells were mounted in the pipe wall to measure the normal tractions, and strain gauges were glued on the inside and outside of the pipe wall to measure the bending moments and thrusts of the pipe. Sand was poured from a hopper into the box with approximately constant height of fall and rate of pouring. To reduce side friction in the box glass plates were used between the sand and steel plate and between the sand and the Perspex plate.

The speed of the centrifuge was then raised in steps of 5g, and the



final load cell readings and strain gauge readings were taken at 35g.

Figure 1 shows the geometry and boundary conditions of the model.

#### IV. FINITE ELEMENT PROGRAM:

The finite element program used in this study, CANDE (Culvert ANalysis and DEsign) was developed by Katona (4,5). Although CANDE can be operated in a design or analysis mode, only the analysis mode was utilized in this study. The main features of CANDE, which are used in the analysis, will be briefly described in this report.

The assumptions common to all solution levels in CANDE are: (1) Two dimensional analysis with plane strain; and (2) small strain theory. There are three solution levels in CANDE. Level 1 corresponds to the elastic, closed solution while levels 2 and 3 are both finite element (FE) solutions. The difference between levels 2 and 3 is in the way the FE mesh is defined. Level 2 generates automatic and symmetric meshes for a circular, elliptical, or box culverts, whereas level 3 allows the user to define a mesh. Since the model culvert used in this study was symmetric and circular, level 2 was used.

There are five material characterizations available for various pipe materials in CANDE. These materials are: (1) steel; (2) aluminum; (3) plastic; and (4) concrete. In this study the concrete material characterization with a bilinear stress-strain relationship was used to represent the concrete material properties.

Three basic elements are contained in the program: a quadrilateral element for representation of the soil, an interface element for slip of elements, and a beam-column element for the culvert.

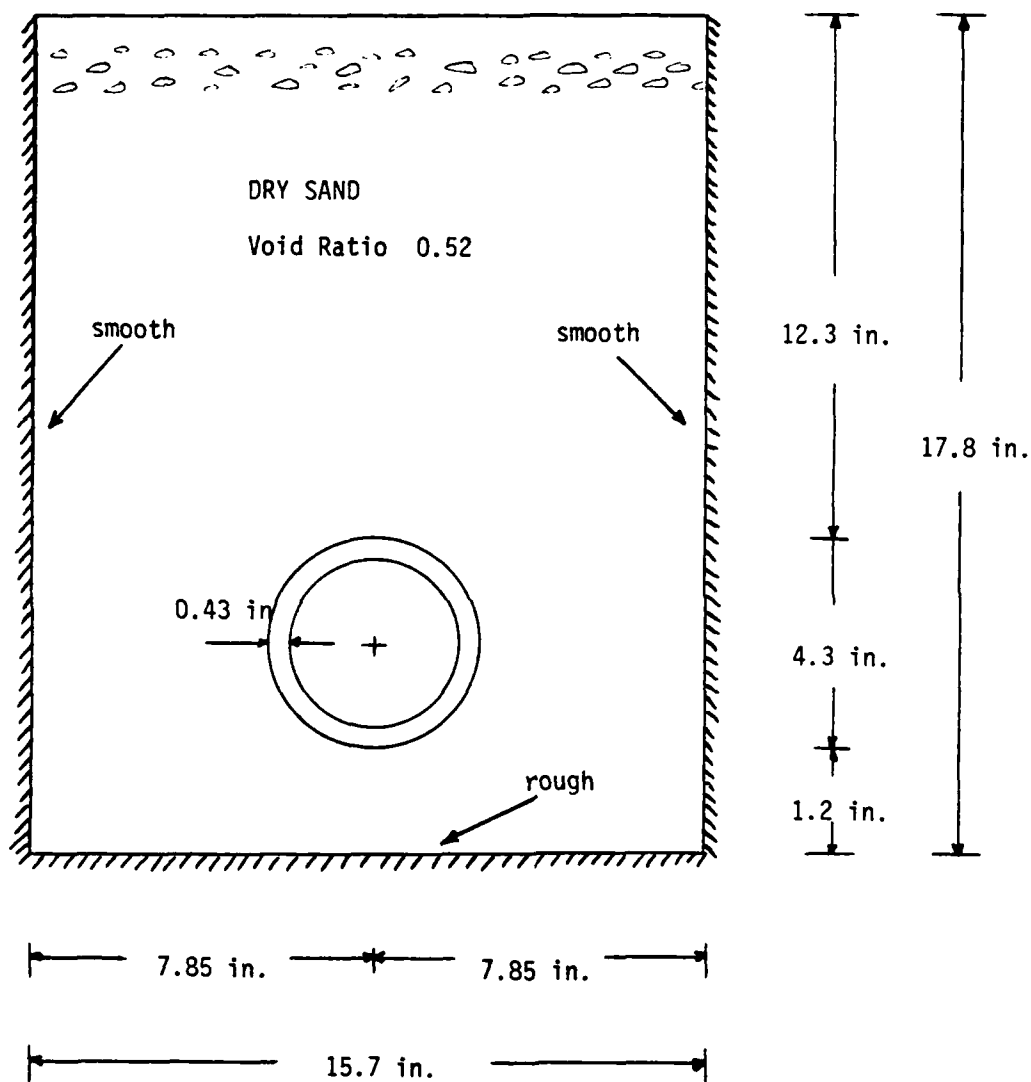


Figure 1. Geometry and Boundry Conditions

For representation of soil behavior, four types of constitutive relationships are incorporated with the program. These are: (1) linear elastic; (2) overburden-dependent; (3) extended Hardin's model; and (4) Duncan's hyperbolic model. In order to investigate the influence of the constitutive relationship for soil on the culvert response in the numerical analysis, both a linear elastic model and a nonlinear model (Duncan's hyperbolic model) for soil were utilized in the analyses. The results for each representation are presented in this report. In reference (3) only dense sand with a void ratio of 0.52 was specified, soil properties for the elastic (FE) soil model were assumed as: Young's modulus of 7,000 psi and Poisson's ratio of 0.3. A "SW" soil type at a compaction of level of 95% in Table 1 was used for Duncan's soil model. These values are within an expected range as recommended by Lambe and Whitman (8) and Duncan (2), respectively.

The finite element grid with boundary conditions used in the comparison analysis is shown in Figure 2. The culvert was represented by 10 beam-column elements, the soil was represented by 86 quadrilateral elements, and the slip model was represented by 11 interface elements.

#### V. RESULTS:

The quasi-theoretical results obtained from the finite element code, CANDE, are compared to the results of the centrifugal testing model. The compared quantities are: (1) normal traction on the culvert periphery; (2) peripheral distribution of bending moment; and (3) peripheral distribution of thrust.

Figure 3 shows graphical comparisons between finite element results

Table 1. Representative Parameter Values of the Modified Duncan Model (Reference 2)

Unified Soil Classification	RC* Stand. AASHTO	$\gamma_m$ k/ft <sup>3</sup>	$\phi_o$ deg	$\Delta\phi$ deg	c k/ft <sup>2</sup>	K	n	$R_f$	$k_b$	m
GM, GP SW, SP	105	0.150	42	9	0	600	0.4	0.7	175	0.2
	100	0.145	39	7	0	450	0.4	0.7	125	0.2
	95	0.140	36	5	0	300	0.4	0.7	75	0.2
	90	0.135	33	3	0	200	0.4	0.7	50	0.2
SM	100	0.135	36	8	0	600	0.25	0.7	450	0.0
	95	0.130	34	6	0	450	0.25	0.7	350	0.0
	90	0.125	32	4	0	300	0.25	0.7	250	0.0
	85	0.120	30	2	0	150	0.25	0.7	150	0.0
SM-SC	100	0.135	33	0	0.5	400	0.6	0.7	200	0.5
	95	0.130	33	0	0.4	200	0.6	0.7	100	0.5
	90	0.125	33	0	0.3	150	0.6	0.7	75	0.5
	85	0.120	33	0	0.2	100	0.6	0.7	50	0.5
CL	100	0.135	30	0	0.4	150	0.45	0.7	140	0.2
	95	0.130	30	0	0.3	120	0.45	0.7	110	0.2
	90	0.125	30	0	0.2	90	0.45	0.7	80	0.2
	85	0.120	30	0	0.1	60	0.45	0.7	50	0.2

\* RC = relative compaction, in percent

Number of Nodes = 132  
Number of Beam Elements = 10  
Number of Quadrilateral Elements = 86  
Number of Interface Elements = 11

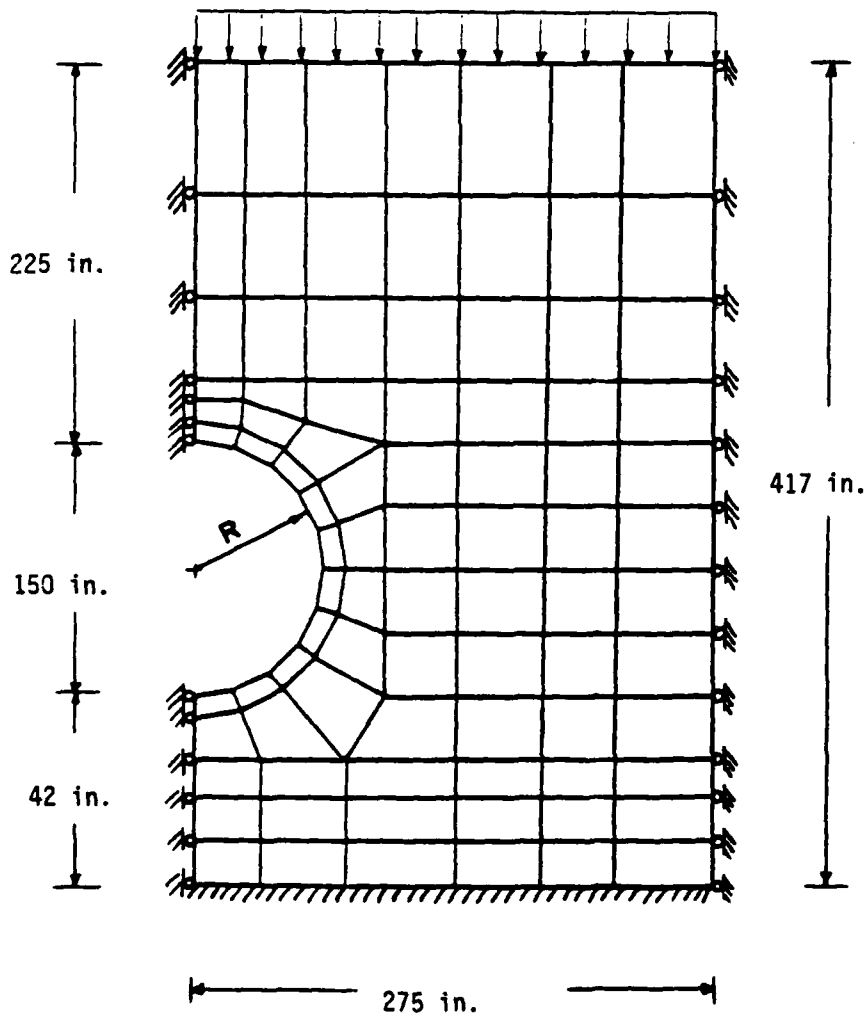


Figure 2. Symmetric Mesh

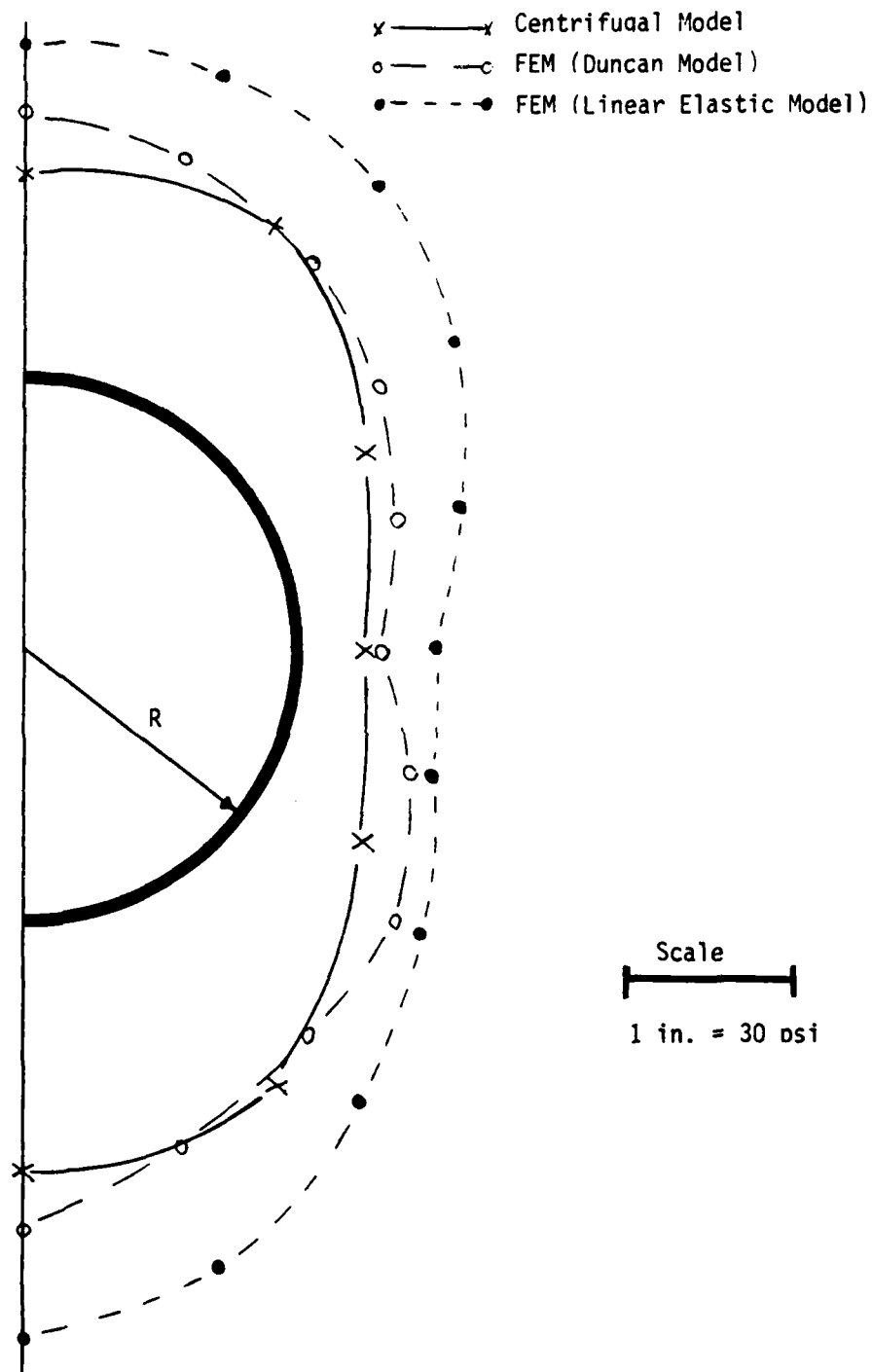


Figure 3. Normal Tractions on the Culvert Periphery

and centrifugal model testing observations for normal tractions acting at the culvert periphery. Excellent comparisons are observed from the results of numerical predictions with Duncan's soil model and the results of the centrifugal model testing. Trends of shape changes are similar and the magnitudes are in excellent agreement. In particular, the magnitudes around the positions of  $30^\circ$ ,  $90^\circ$  (springing line),  $144^\circ$ , and  $162^\circ$  from the crown are almost identical each other.

Yet, discrepancies exist between the numerical predictions with the linear soil model and other results obtained from the centrifugal testing model and numerical predictions with nonlinear soil model - even though the trend of shape change is similar.

Figure 4 shows internal, peripheral bending moment distributions of the culvert. As can be seen from the figure, moderately good comparisons are observed between the results of numerical analyses and the results of the centrifugal testing model. Shape changes are similar and the magnitudes are in good agreement except around the invert. Positive moment in the figure is that which produces tension at the inner fibers of the concrete.

Figure 5 compares the results of finite element analyses and the centrifugal model study of circumferential thrusts of the culvert. Numerical predictions and model measurements are similar in shape, but do not agree well in magnitudes around the crown. Correlations are very good around the springing line and the invert.

Again, the results of CANDE with an elastic soil model are somewhat conservative when compared to other results.

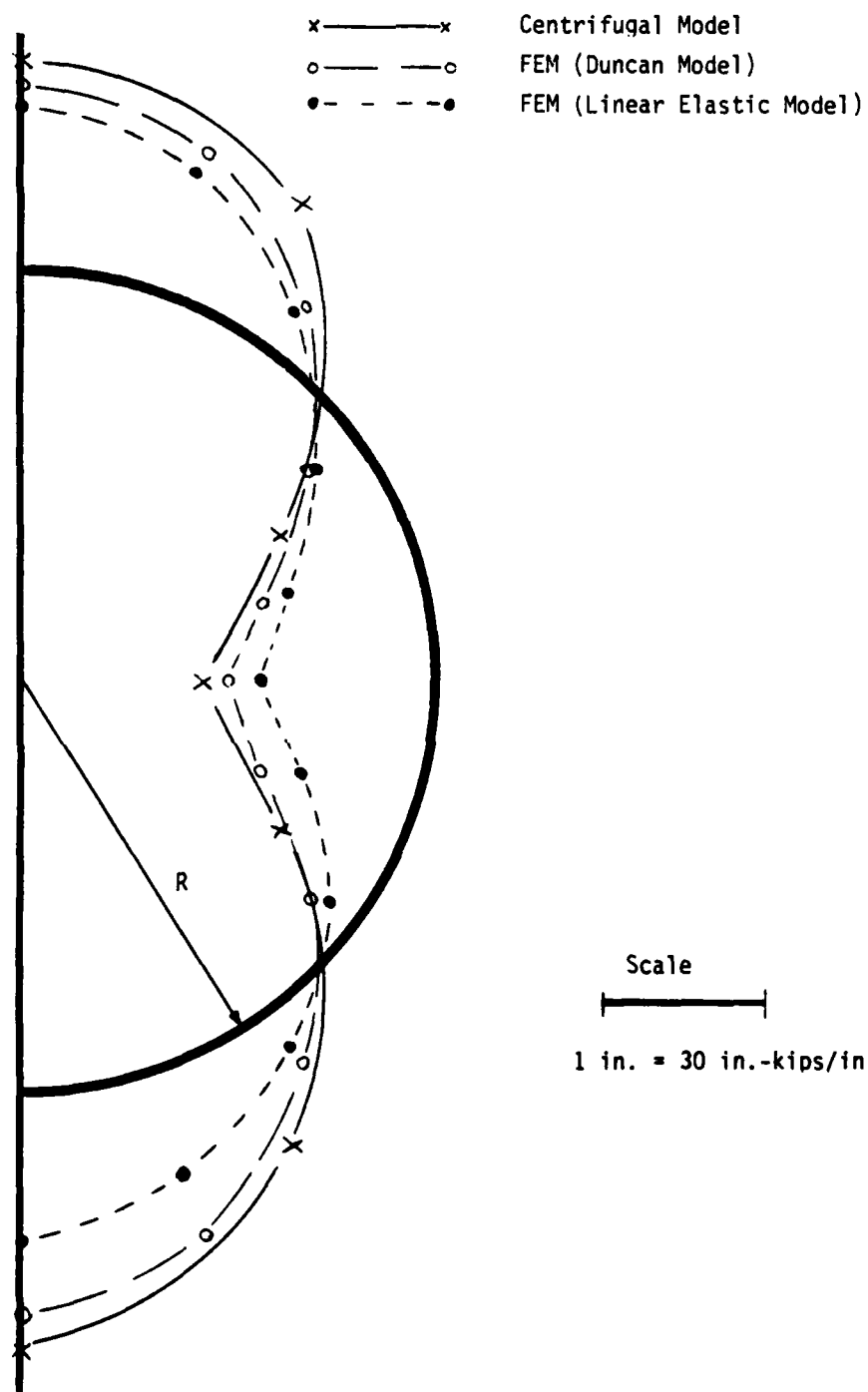


Figure 4. Peripheral Distribution of Bending Moment



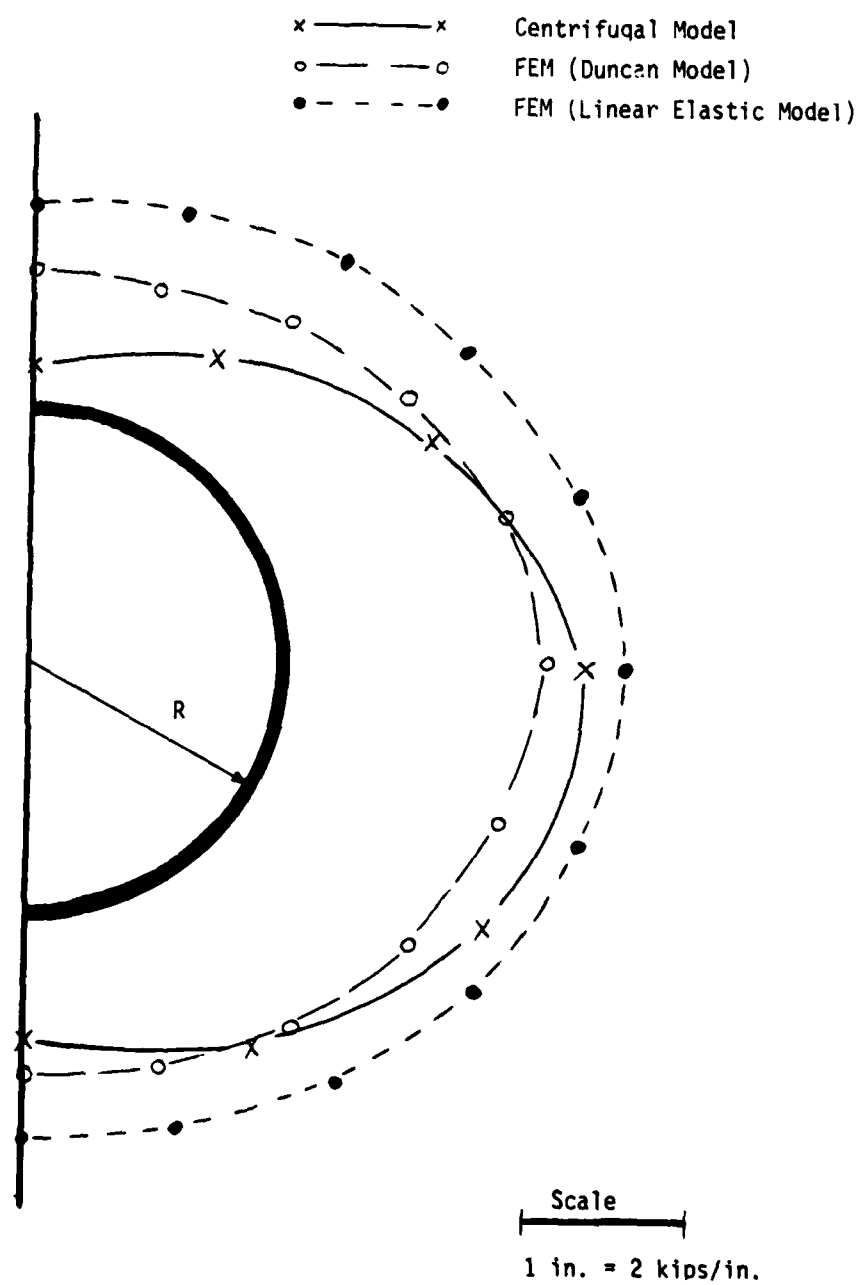


Figure 5. Peripheral Distribution of Thrust

## VI. RECOMMENDATIONS:

This study evaluates the capabilities of a finite element method for analyzing the performance of buried structures by comparing the predicted and measured behavior of a concrete culvert. The findings and recommendations for further research based on this study are as follows:

1. The results of the finite element analyses compared well in shape with the results of the centrifugal testing model. Analysis with Duncan's hyperbolic relationships for soil exhibits a more realistic approach for predicting soil and culvert responses; the results were much closer to the results of the centrifugal testing model in shapes and magnitudes.
2. Significant economic savings (time and costs of construction and operations) could be obtained by using the centrifugal modeling technique to study the behavior of prototype buried culverts if results of the centrifuge model testing agree well with field measurements.
3. Additional economic savings could be obtained when numerical (finite element) analyses are used with realistic constitutive models for soils and culverts.
4. Additional studies (i.e. box, arch, or elliptical culverts with different bedding and boundary conditions) should be conducted for further verifications before general acceptance can be awarded.

#### REFERENCES

1. Costes, N.C., and Proudley, C.E., "Performance Study of Corrugated-Metal-Pipe Culvert Under Embankment-North Carolina," Flexible Culverts Under High Fills: Highway Research Board Bulletin 125, National Academy of Science, Washington, D.C., 1956, pp. 58-169.
2. Duncan, J.M., et al., "Strength, Stress-Strain and Bulk Modulus Parameters for Finite Element Analyses of Stresses and Movements in Soil Masses," Report No. UCB/GT/78-02, National Science Foundation, April, 1978.
3. James, R.G., and Larsen, H., "Centrifugal Model Tests of Buried Rigid Pipes," Proceedings of the Ninth International Conference on Soil Mechanics and Foundation Engineering, Tokyo, Japan, 1977.
4. Katona, M.G., Smith, J.M., Odello, R.J., and Allgood, J.R., "CANDE - A Modern Approach for the Structural Design and Analysis of Buried Culverts," Report No. FHWA/RD-77/5, Federal Highway Administration, Washington, D.C., October, 1976.
5. Katona, M.G., Vittes, P.D., Lee, C.H., and Ho, H.T., "CANDE-1980: Box Culverts and Soil Models," Report No. FHWA/RD-80/172, Federal Highway Administration, Washington, D.C., May, 1981.
6. Kim, Y.S., and Davis, R.E., "Proof Testing of a Structural Plate Pipe with Varying Bedding and Backfill Parameters Section VII, Volume 3: Finite Element Analysis," Report No. FHWA/CA/SD-83/04, California Department of Transportation, Sacramento, California, August, 1983.
7. Lambe, T.W., and Whitman, R.V., Soil Mechanics, John Wiley and Sons, Inc., New York, N.Y., 1969.
8. Leonards, G.A., Wu, Tzong-Hsin, and Juang Charng-Hsein, "Predicting Performance of Buried Conduits," Report No. FHWA/IN/JHRP-81/3, Indiana State Highway Commission, Indianapolis, Indiana, June, 1982.
9. Marston, A., and Anderson, A.O., "The Theory of Loads on Pipes and Ditches," Bulletin 31, Iowa Engineering Experimental Station, Ames, Iowa, 1913.
10. Spangler, M.G., "Underground Conduits: An Appraisal of Modern Research," Transactions, ASCE, Vol. 113, June, 1947, pp. 316-374.
11. Rowe, P.W., "Large Scale Laboratory Model Retaining Wall Apparatus," Proceedings of Roscoe Memorial Symposium, Cambridge, England, 1971, pp. 279-289.

1984 USAF-SCEEE Summer Faculty Research Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Southeastern Center for Electrical Engineering Education

FINAL REPORT

Acoustic Emission in Composites

Prepared by: Ronald A. Kline/Kevin M. White

Academic Rank: Assistant Professor/Graduate Student

Department and University: School of Aerospace, Mechanical  
and Nuclear Engineering  
University of Oklahoma

Research Location: Air Force Materials Laboratory  
Nonmetallic Materials Division

USAF Research: Dr. Stephen Tsai

Date: August 10, 1984

Contract No.: F49620-82-C-0035

# Acoustic Emission in Composites

by  
Ronald A. Kline

## ABSTRACT

Acoustic emission (AE) phenomena in composite materials is examined both theoretically and experimentally. Theoretically, the effect of material attenuation is introduced into an AE model to aid in separating source characteristics from propagation effects in measured AE signals. AE waveforms in composite samples were measured for controlled experiments. The source for these experiments was primarily restricted to the fracture of a single graphite fiber in an epoxy matrix and localized at a single point. The normal component of the surface displacement was measured at epicenter using a point contact, conical sensor. In contrast to previous work in this field, in analyzing the signals attention was focused on the initial portion of the AE waveform rather than long term signal behavior which is dominated by internal specimen reflections. For comparison, AE waveforms on fiber free samples are also shown.

#### Acknowledgement

The authors wish to thank the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to conduct their research at the Air Force Materials Laboratory, Wright-Patterson Air Force Base, Ohio. In addition, we wish to acknowledge the assistance of personnel of the Nonmetallic Materials Division, Mechanics and Surface Interaction Branch and the University of Dayton Research Institute in this research effort. Special thanks go to Dr. L. Drzal (AFML) and M. Rich (UDRI) for their help in the design and fabrication of test specimens. The authors would also like to thank Dr. S. Tsai (AFML) for his support and encouragement throughout the course of this program.

## Introduction

Acoustic emission (AE) techniques have proven to be extremely useful in both materials research and nondestructive testing applications. However, much remains to be accomplished before AE fulfills its complete promise as an investigative method. Many of the details of acoustic emission signal generation are poorly understood despite the large amount of effort devoted to this end. Thus, the vast majority of acoustic emission applications are qualitative rather than quantitative in nature. This is particularly true of composites. Most of what we know about AE behavior in composites consists of general correlations between observations of mechanical behavior and AE. Ideally, one wishes to be able to go beyond these simple correlations.

## Objectives

Given the relative complexity of the composite microstructure and the multitude of possible deformation mechanisms, one might wish to separate out AE from different sources to better assess their potential impact on the load bearing capability of the test piece. Potentially harmful AE sources as well as benign sources should each have a characteristic signature which would allow such a classification. The development of such a classification scheme is the ultimate objective of this research program. Obviously, this is a rather ambitious goal given the present state of state of the art and the short duration of this summer program. However, by concentrating our efforts on two aspects of the problem we were able to make some progress toward this end.

One area of particular interest is the precise nature of the interaction of the acoustic waves associated with AE and the propagation medium. While there has been considerable research devoted to developing new models of AE

sources to more realistically reflect actual generation mechanisms<sup>(1,2)</sup>, material attenuation has been overlooked in the development of these models. This is despite the fact that attenuation can have a profound influence on the characteristics of the observed AE signals (rise time, amplitude, etc.). As many applications of acoustic emission take advantage of the remote sensing capabilities of the method, this factor cannot be neglected. Attenuation is also of importance when dealing with highly attenuating media such as polymers and polymer based composites. The second area of interest is the study of acoustic emission from specific sources of importance in composite testing. In order to identify the characteristic features of the various types of acoustic emission which can occur, each mechanism must be isolated and studied independently. In this investigation only one mechanism (fiber fracture) is considered. Ultimately it is hoped that the research program will be expanded to include a wide variety of AE sources and form the basis for a comprehensive signature library.



## Part I: The Attenuation of Acoustic Emission Signals

### Background

An acoustic emission signal can be represented as being composed of three distinct parts in the form

$$u(t) = S(t) * M(t) * T(t) \quad (1)$$

where  $S(t)$  represents the contribution of the AE source itself,  $M(t)$  represents the interaction of the acoustic emission waveform with the propagation media (both via interaction with the specimen boundaries as well as attenuation), and  $T(t)$  represents the effect of the sensor being used. In much of the early work in acoustic emission, sensor characteristics were a limiting factor in the interpretation of acoustic emission signals. This problem arose from the use of conventional piezoelectric sensors which, although sensitive, introduced considerable changes in the observed AE waveform. Among the problems associated with these sensors has been their large surface area (so one obtains an integrated measure of the transducer response rather than response at a point), the fact that they are contact sensors (and can distort the waveform) and the inability to relate their response to any single physically meaningful parameter (although Hsu et. al.<sup>(3)</sup> report that the response more closely resembles displacement than anything else). Once this was realized, alternate sensors were developed to overcome many of these limitations notably the capacitance sensor<sup>(4-6)</sup>, the electromagnetic sensor<sup>(7)</sup> and the optical sensor<sup>(8)</sup>. While less sensitive than conventional piezoelectric elements, the new sensors have been shown to be easily calibrated, point displacement sensors and have proven their value in numerous studies of acoustic emission phenomena, principally with simulated AE sources<sup>(3,5-6)</sup> or highly energetic events<sup>(9-11)</sup>. For more realistic AE

amplitudes, efforts have been focussed on determining sensor characteristics and mathematically removing their effects from the observed waveforms using signal deconvolution techniques (the so-called transfer matrix approach<sup>(12)</sup>). Alternatively, one can use a hybrid piezoelectric sensor developed by Proctor<sup>(13)</sup> which employs a conical piezoelectric element, rather than a flat plate, to yield point displacement results without sacrificing the sensitivity and operating convenience of piezoelectrics.

A second factor which influences AE signal content is the effect of specimen geometry and composition. When an acoustic wave strikes a boundary, reflection and mode conversion can occur. Repeated reflection can introduce resonances into the AE signal. This factor has made the frequency analysis of acoustic emission signals an extremely difficult process. In order to avoid this problem, investigators have taken great pains to localize the AE source and utilize specimens which minimized internal reflections. This type of an approach was employed by Breckenridge et. al<sup>(5)</sup> in their watershed research which conclusively established the fundamental relationship between elastodynamic theory and acoustic emission. Since this employed an artificial, rather than actual AE source, it was relatively easy to achieve an approximation to a reflection free medium with a localized source. For actual AE sources, the task has been more difficult. Scruby and Wadley<sup>(10)</sup> have developed an ingenious way of eliminating the reflection problem to a large extent and localizing the use of a narrow gage test section with conical grip areas which they refer to as a 'yobel'. Material attenuation is another way in which the specimen can influence the AE waveform. This can occur via a polycrystalline media and hysteresis losses in viscoelastic media. The obvious goal of AE research has been to eliminate the experimentally induced effects in order to focus on source characteristics. The signal deconvolution methods

discussed earlier can be used. At least in principle, to eliminate attenuation effects as well as transducer characteristics in the signal analysis.

The final factor to be considered is the nature of the source itself. Most of the work in this area was based upon the theory developed to explain seismic behavior<sup>(1)</sup>. Most of this work has involved idealized sources and geometries such as that of Lamb's problem of a point force on the surface of a semi-infinite half space. In experiments with sources and geometries designed to approximate the theoretical models eg. the fracture of a glass capillary tube by a point indenter to represent the point force and a large block. The results have been in agreement with the theoretical predictions. Correlation of these results with actual AE sources has been somewhat less successful. This lack of agreement has lead to the development of more complicated models to more realistically represent AE generation. Pao and co-workers<sup>(2,14-15)</sup> have been particularly active in this area with the introduction of alternatives to the point force (including the double force, coupled forces and the center of dilatation) as well as obtaining results for off axis response away from epicenter. Several investigators have modified the time behavior of the source from the heaviside step function behavior assumed in the elementary theories to a more realistic ramp function to better reflect the rise time of the source. Comparison of the results from the newer theories has been encouraging<sup>(16)</sup> but more work remains in this area before this connection is conclusively established.

#### Theory

In this section the effects of the propagation medium on acoustic emission waveforms are discussed. The technique described here relies on a

Fourier decomposition of an AE signal into its spectral components. The geometry is illustrated in Figure 1. The source is taken to be a ramp function

$$f(t') = \begin{cases} 0 & t' < 0 \\ F_0 \frac{t'}{\tau} & 0 < t' < \tau \\ F_0 & t' > \tau \end{cases} \quad (2)$$

The origin of the coordinate system is located at the source. For simplicity only the displacement  $u$  at epicenter is considered. To further simplify matters, attention is restricted to the normal component of the surface displacement  $U_z$  as this is the component which can be readily measured. The in-plane components can be easily obtained using the approach outlined here (with a suitable modification of the Green's function).

The surface displacement at the source is then given by the convolution of the forcing function at the source with the appropriate Green's function. In this research, the approach of Shibata<sup>(16)</sup> is utilized for the Green's function. Shibata employed the Green's function for an infinite space  $G_{zz}^{\infty}$  for the AE source at (0,0,0) along with an image force at (0,0,2R) to compensate for the effects of the free surface. This simplifies the mathematical representation for the Green's function with a minimal effect on the results. For this case the Green's function  $G_{33}^{\infty}$  is given by

$$G_{33}^{\infty}(R, \bar{t}) = \frac{1}{4\pi e} \left( \frac{1}{A^2 R} \delta(\bar{t} - \frac{R}{A}) - \frac{2}{R^2} \bar{t} H(\bar{t} - \frac{R}{A}) - H(\bar{t} - \frac{R}{C}) \right)$$

where

$\rho$  = density

$A$  = longitudinal wave speed

$c$  = shear wave speed

$\bar{t} = t - t'$

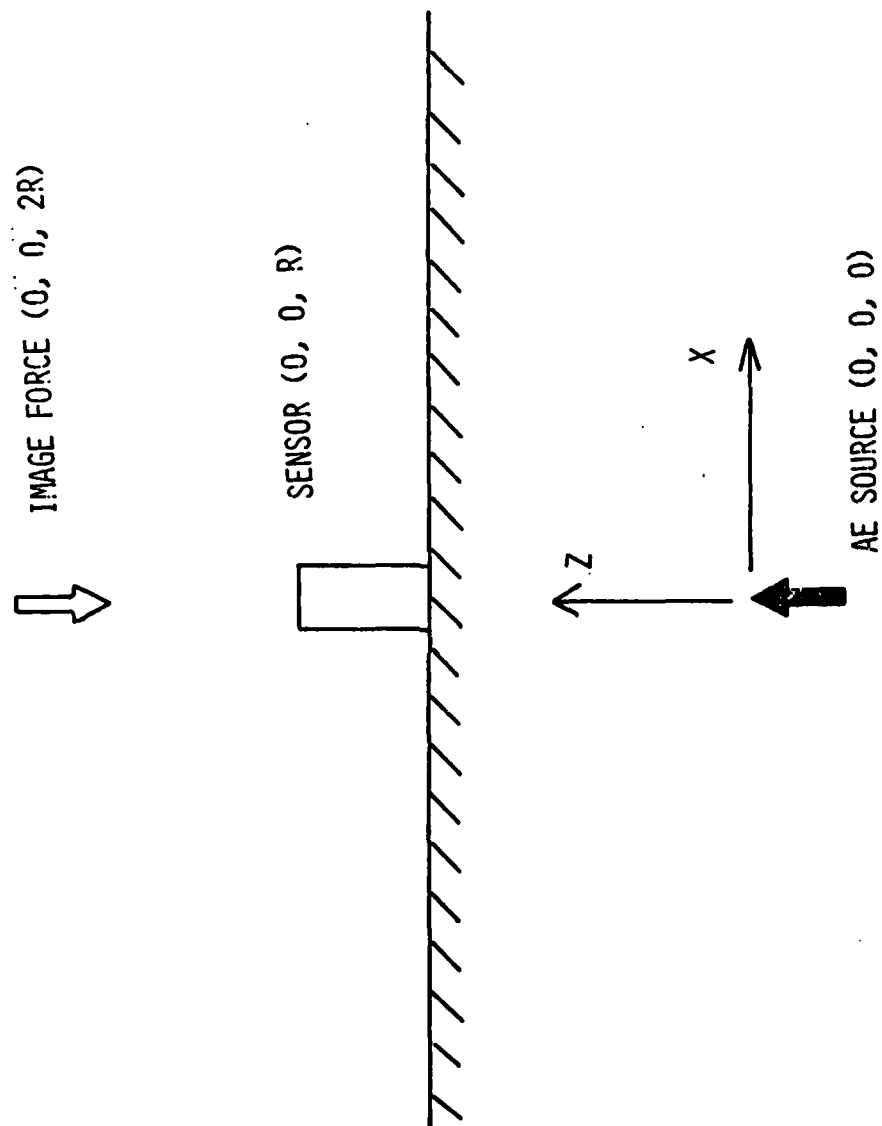


FIGURE 1. AE GEOMETRY

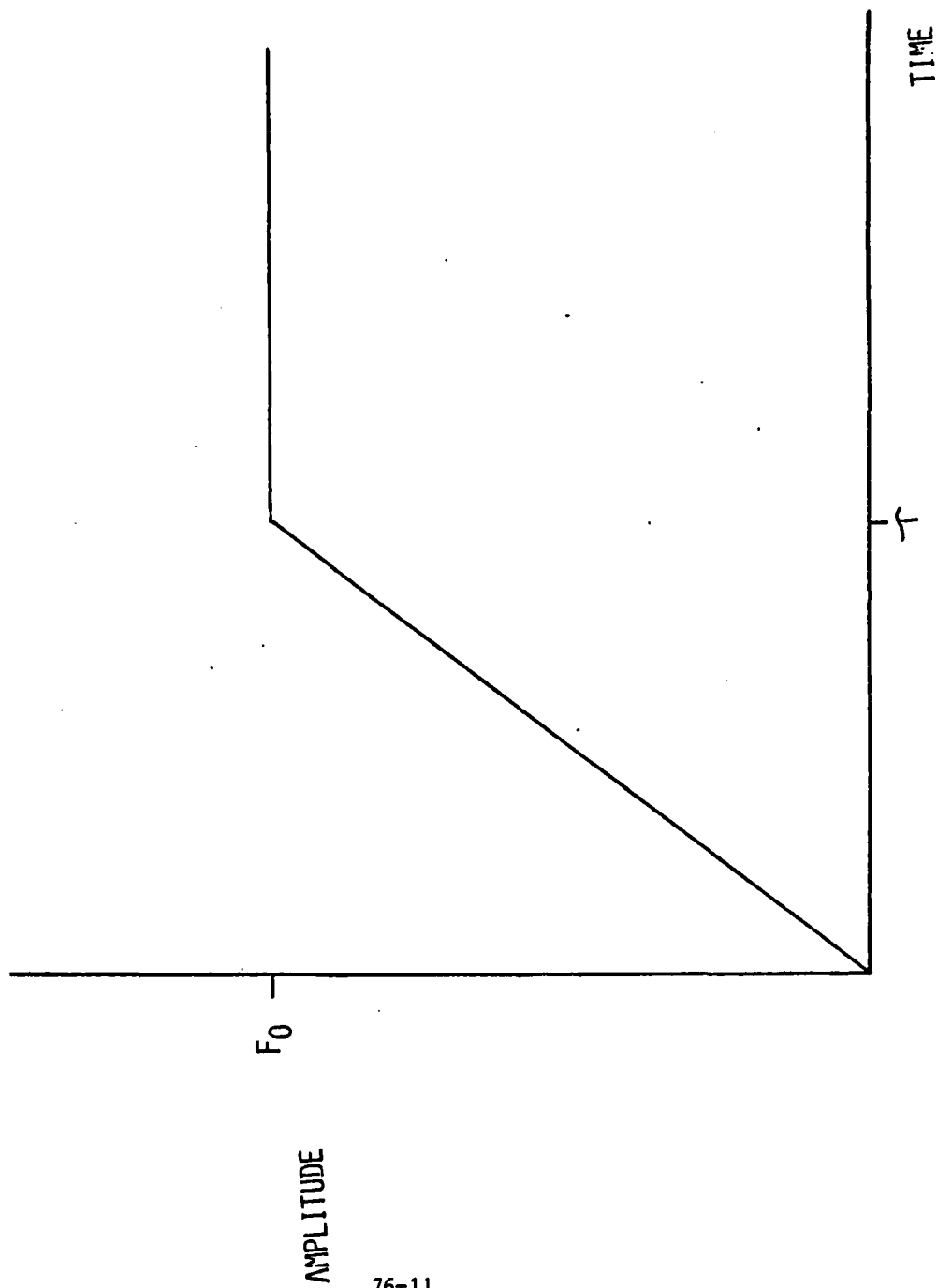


FIGURE 2. AE SOURCE - RAMP FUNCTION

It is readily apparent that the Green's function is composed of elements associated with the arrival of longitudinal waves and shear waves. Since these components are attenuated at different rates, it is advantageous to separate the Green's function into two parts,  $G$  and  $G_s$ , associated with the two types of waves. The surface displacement is then found to be

$$u_z(R,t) = U_z^L(R,t) + U_z^S(R,t)$$

where

$$\begin{aligned} U_z^L(R,t) &= 2V \int_0^t G_L^{\infty}(R,\bar{t}) f(t') dt' \\ U_z^S(R,t) &= 2V \int_0^t G_s^{\infty}(R,\bar{t}) f(t') dt' \\ G_L^{\infty}(R,t) &= \frac{-1}{4\pi\rho} \left[ \frac{1}{A^2 R} \delta\left(t - \frac{R}{A}\right) + \frac{2}{R^2} t H\left(t - \frac{R}{A}\right) \right] \\ G_s^{\infty}(R,t) &= \frac{-1}{4\pi\rho} \left[ -\frac{2}{R^2} \bar{t} H\left(t - \frac{R}{C}\right) \right] \end{aligned} \quad (3)$$

This is illustrated in Figure 3 for the source function described in Equation 2 (and no attenuation) with  $U_z^L$ ,  $U_z^S$  and  $U$  being presented. From this figure it is apparent that the principal effect of the shear component is to negate the contribution of the longitudinal component for all times after the shear arrival ( $T = R/C$ ). For the case where the source is a ramp function and attenuation is absent, the displacement terms  $U_z^L$  and  $U_z^S$  can be readily calculated. This is presented in Appendix 1.

In introducing the effects of attenuation into this model, the displacements are considered to be formed by the superposition of their various spectral components. Taking the Fourier transform of the two displacements, we have

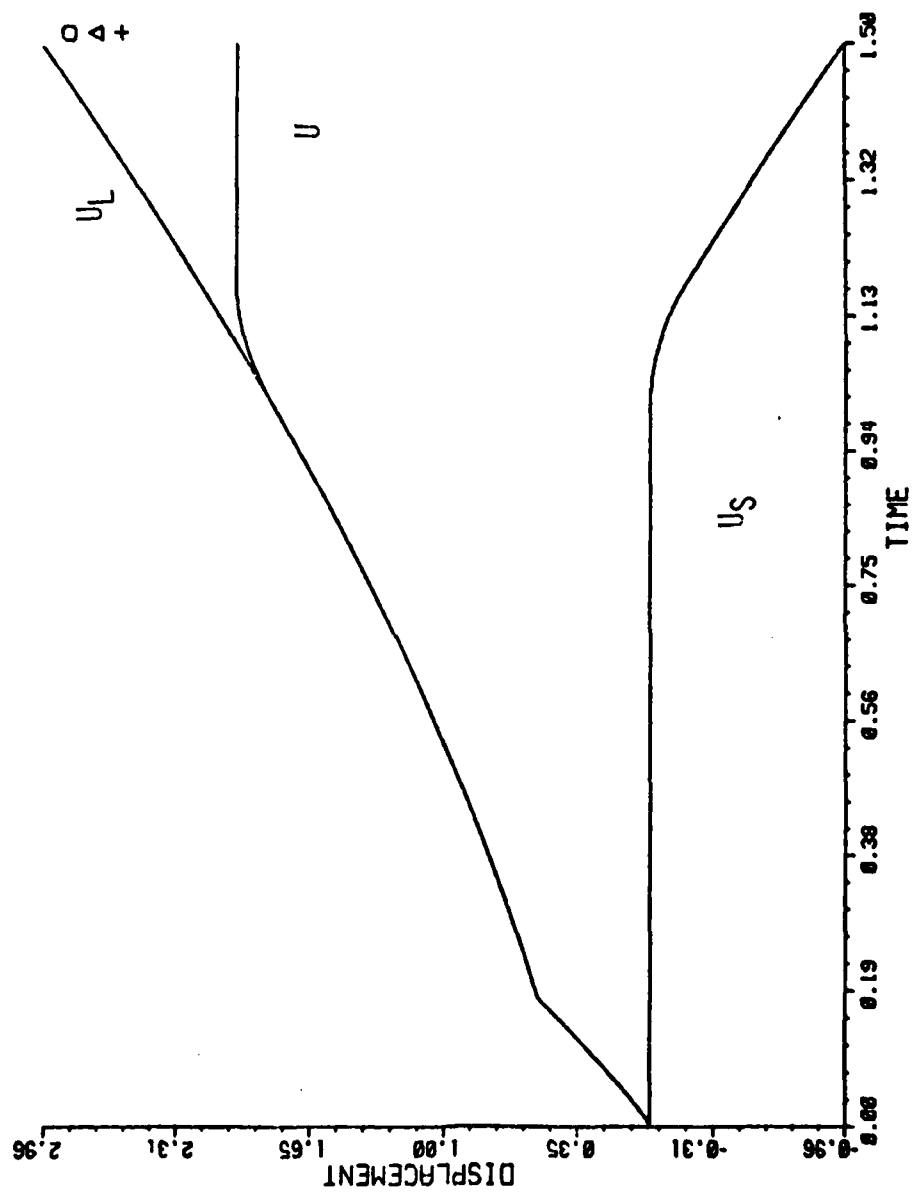


FIGURE 3.  $U$ ,  $U_L$ ,  $U_S$ , VS TIME



$$\bar{U}_Z^L(R, \omega) = \int_{-\infty}^{\infty} U_Z^L(R, t) e^{-i\omega t} dt = M^L(R, \omega) e^{i\phi^L(R, \omega)}$$

$$\bar{U}_Z^S(R, \omega) = \int_{-\infty}^{\infty} U_Z^S(R, t) e^{-i\omega t} dt = M^S(R, \omega) e^{i\phi^S(R, \omega)}$$

where

- Denotes the Fourier transform of the quantity

$M(R, \omega)$  = Amplitude spectrum of the waveform

$\phi(R, \omega)$  = Phase spectrum of the waveform

and L, S indicate the type of wave

The principal effect of attenuation is that it will modify the amplitude spectrum of the waveform. If dispersion was present a suitable modification would have to be made in the phase spectrum. For the materials considered here the dispersion is sufficiently small to be safely neglected. Since attenuation is known to be highly frequency dependent this must be done using an appropriate attenuation factor for each frequency component present in the signal. If we consider plane wave propagation of the form

$$\underline{u}(x, t) = \underline{A}_0 e^{-\alpha x} e^{i(Bx - \omega t)}$$

The effect of attenuation is represented by the decay constant  $\alpha$ . For polychromatic signals the representation is similar with a taking on a frequency dependent form  $\alpha(\omega)$ . For polycrystalline materials  $\alpha$  can be represented as<sup>(17)</sup>

$$\alpha(\omega) = K_1 \omega^4 \quad (4)$$

while polymers are better represented with<sup>(18)</sup>

$$\alpha(\omega) = K_2 \omega \quad (5)$$

It should be noted that shear and longitudinal wave propagation are associated with different attenuation constants in these models. Once the amplitude

spectra are modified to account for attenuation the surface displacements can be directly evaluated using the inverse Fourier transform as shown. Thus we have

$$U_z^L(R,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\alpha_L(\omega)R} M^L(R,\omega) e^{i\phi^L(R,\omega)} e^{i\omega t} d\omega$$

$$U_z^S(R,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\alpha_S(\omega)R} M^S(R,\omega) e^{i\phi^S(R,\omega)} e^{i\omega t} d\omega$$

The computer algorithm used for these calculations is presented in Appendix 2.

#### Experimental Procedure

In order to test the validity of this approach, a brief experimental program was conducted. Specimens of polymethyl methacrylate (PMMA) of three different thicknesses (0.30 cm., 0.59 cm., and 0.88 cm.) were tested. A specially designed piezoelectric transducer with a conical element was employed. This sensor has been found to be a good approximation to the true point displacement sensors with a relatively flat frequency response over the bandwidth of interest and enhanced sensitivity over the capacitive and optical sensors. A Nicolet 204-A digital oscilloscope was used to capture the AE events. This scope had a maximum sample rate of 50 nSEC/point which, using Nyquist's sampling theorem, translates to a frequency resolution of 10 MHz.

The fracture of pencil lead of two different thicknesses (0.3 mm. and 0.5 mm.) was used to simulate an AE source. The transducer was placed at epicenter and the waveforms recorded. Permanent records of each waveform were obtained from the analog scope output using an X-Y recorder. The experimental configuration is shown in Figure 4.

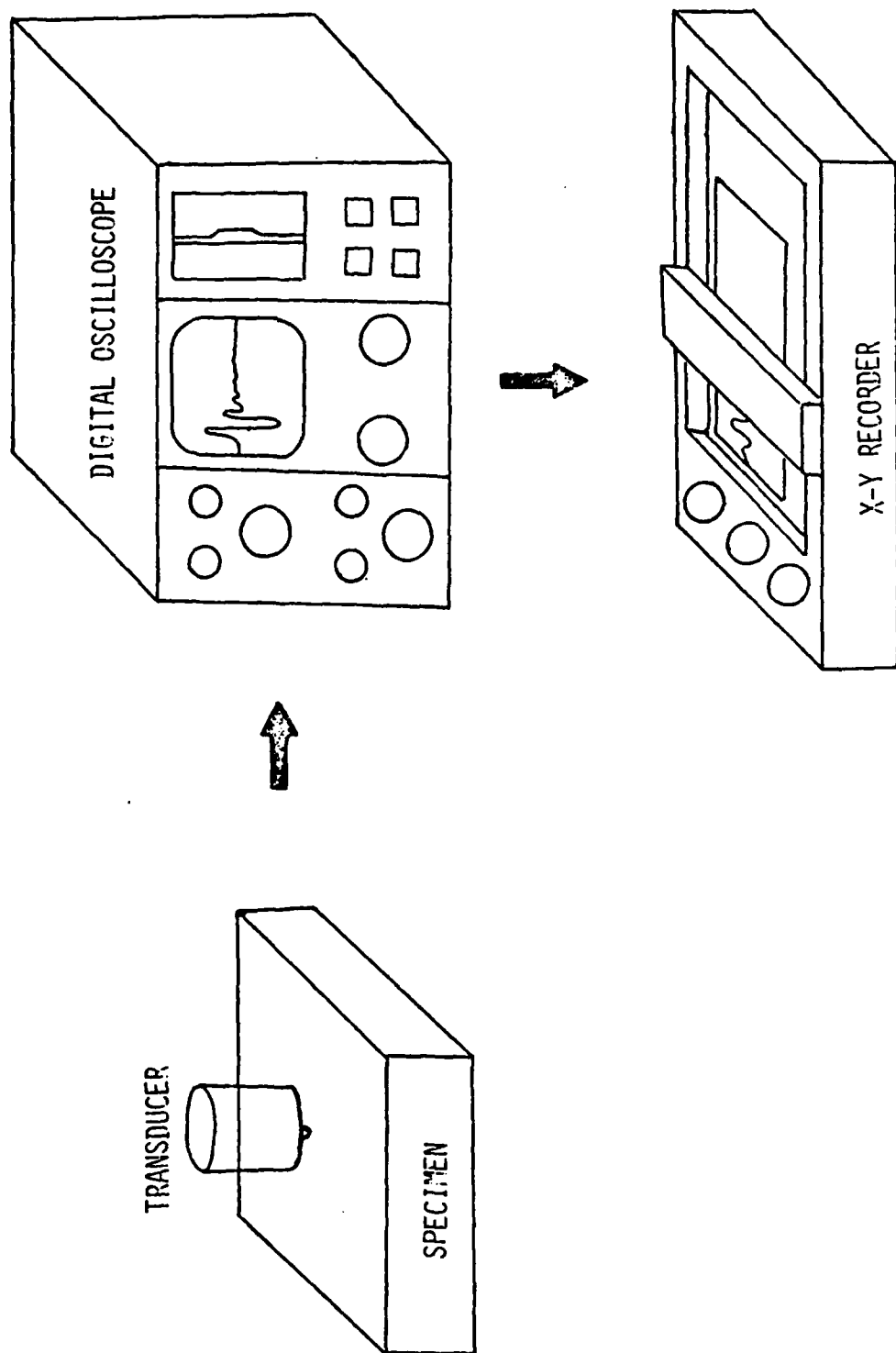


FIGURE 4. EXPERIMENTAL CONFIGURATION

## Results and Discussion

In order to examine the effects of attenuation, let us first consider the response of two aluminum samples (thickness = 1.0 cm., 2.5 cm.) to the same AE event (Figure 5). The AE source is assumed to be of unit amplitude with a rise time of 300 NS. The principal differences between the two waveforms are due to the increased time delay between the longitudinal ( $T = R/A$ ) and shear ( $T = R/C$ ) arrivals in the thicker specimen as well as a reduction in amplitude. Upon further examination it is apparent that the bulk of this reduction is due to a geometric factor (beam spread). In Figures 6 and 7 the results for a 2.5 cm. thick sample are compared with the results for a relatively large propagation distance of 25.0 cm. Waveforms are shown both with and without attenuation to illustrate the magnitude of the reduction attributable to material attenuation rather than beam spread. To facilitate this comparison the time axes have been normalized by the longitudinal ( $T = 0.0$ ) and shear ( $T = 1.0$ ) arrivals. For the 2.5 cm. thick sample one observes a slightly slower signal rise for the attenuated signal, approximately 10 percent slower than that predicted for the same source behavior without attenuation in the model. For the thick (25.0 cm.) sample much the same behavior is observed with a 20 percent change in the rise observed. It should be noted that in order to compare the signals it was necessary to subtract an offset value from the attenuated signal so that the displacement was zero prior to the arrival of the longitudinal component. This constant offset superimposed on the signal appears to be an artifact of the Fourier transformation process.

Calculations were also carried out for two polymeric media, PMMA and PE. These materials were chosen because of the wealth of experimental data available on ultrasonic attenuation in these materials. Figures 8-11 illustrate the larger effect that attenuation has in polymers for the same

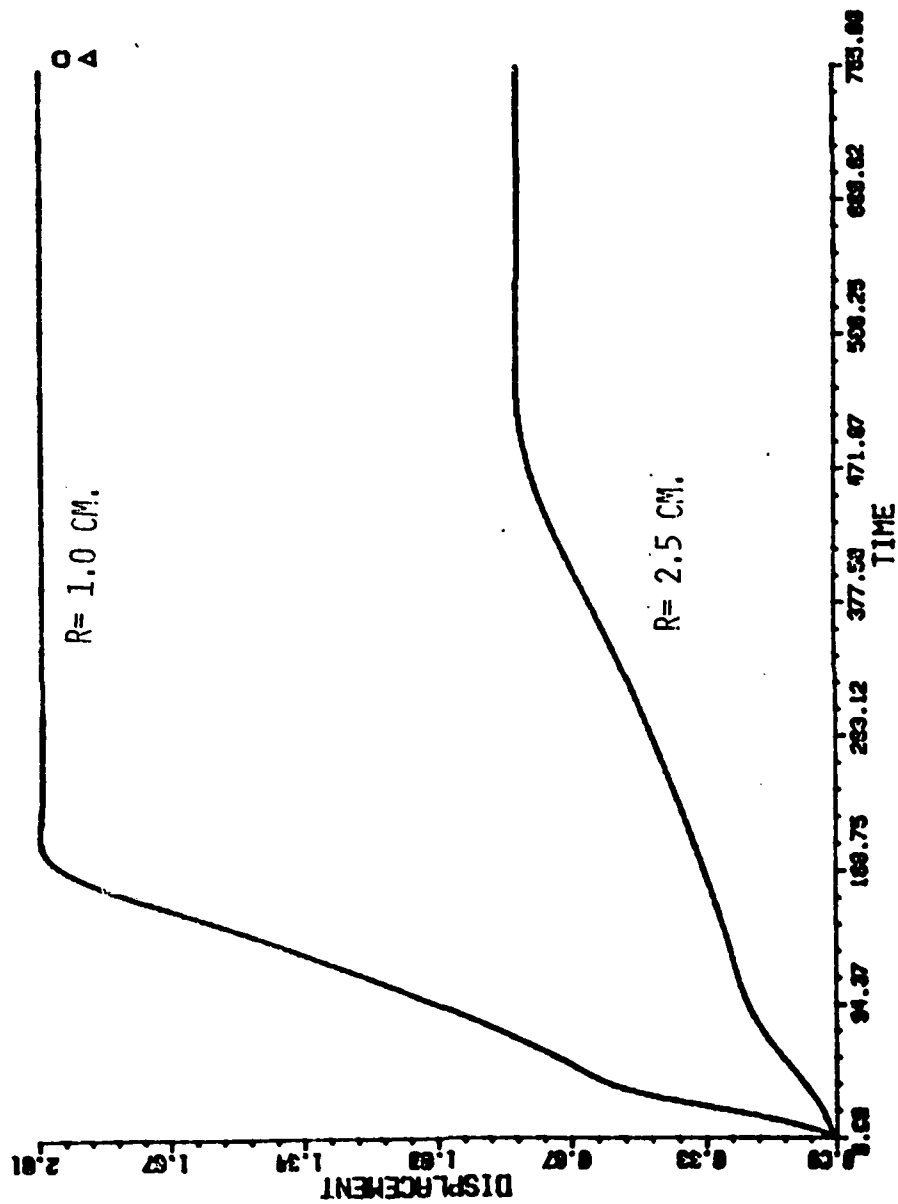


FIGURE 5. ALUMINUM

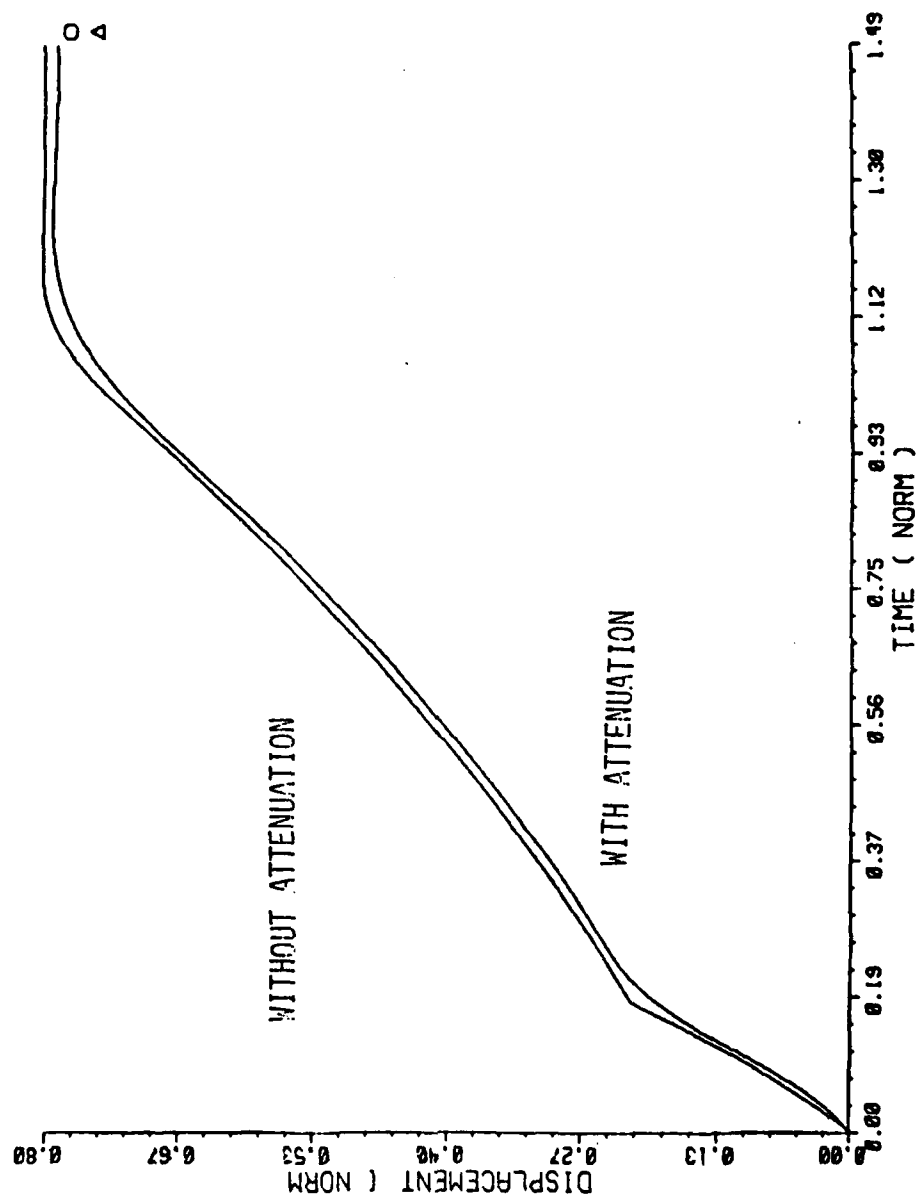


FIGURE 6. ALUMINUM (R= 2.5 CM.)

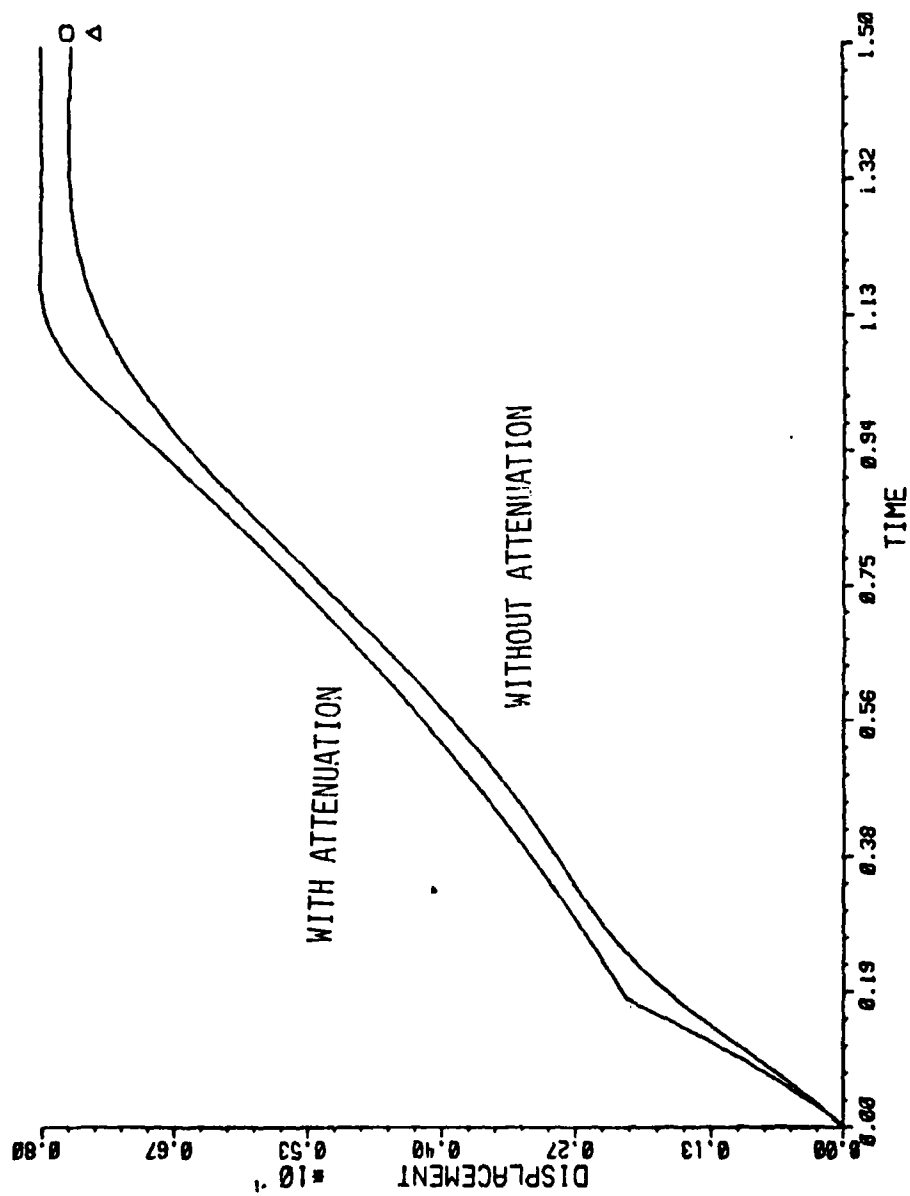


FIGURE 7. ALUMINUM (25.0 CM.)

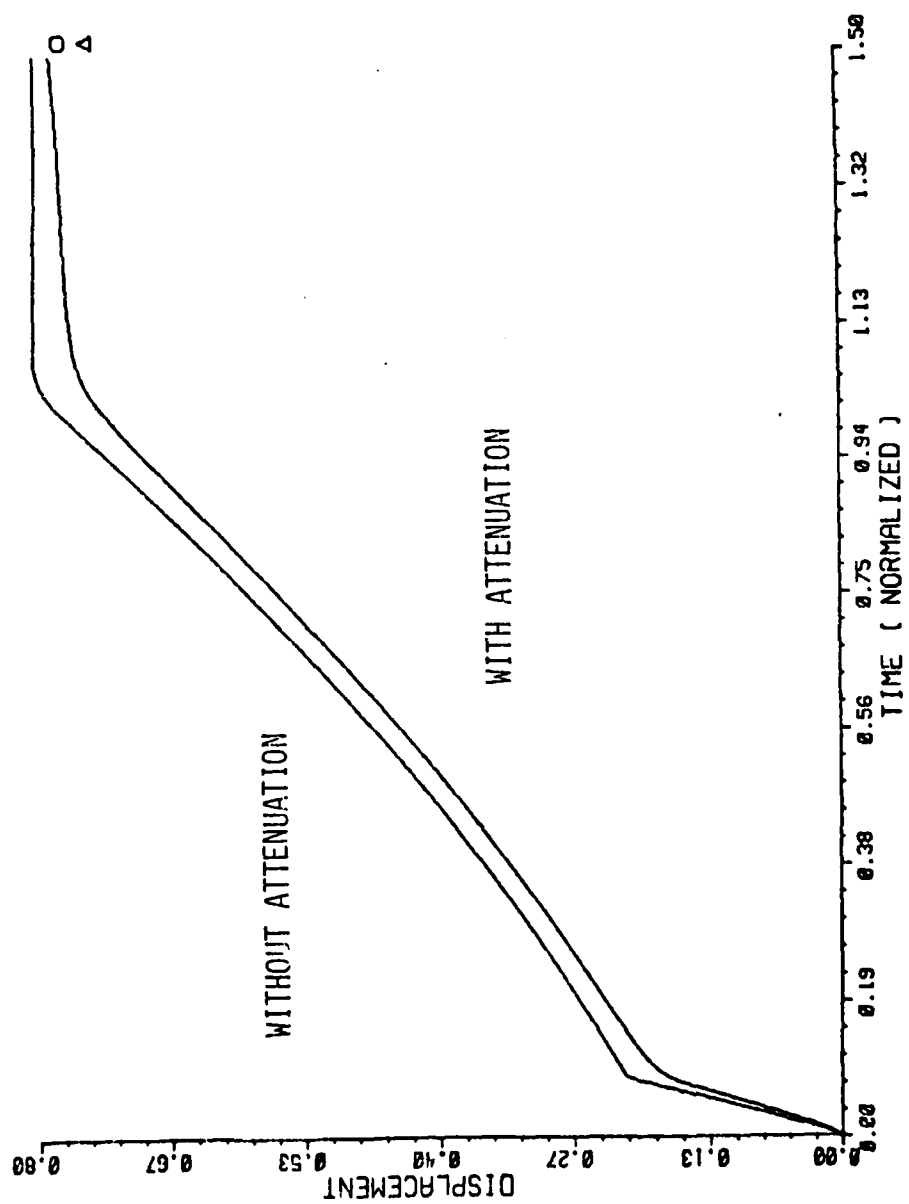


FIGURE 8. PMMA (R= 2.5 CM.)



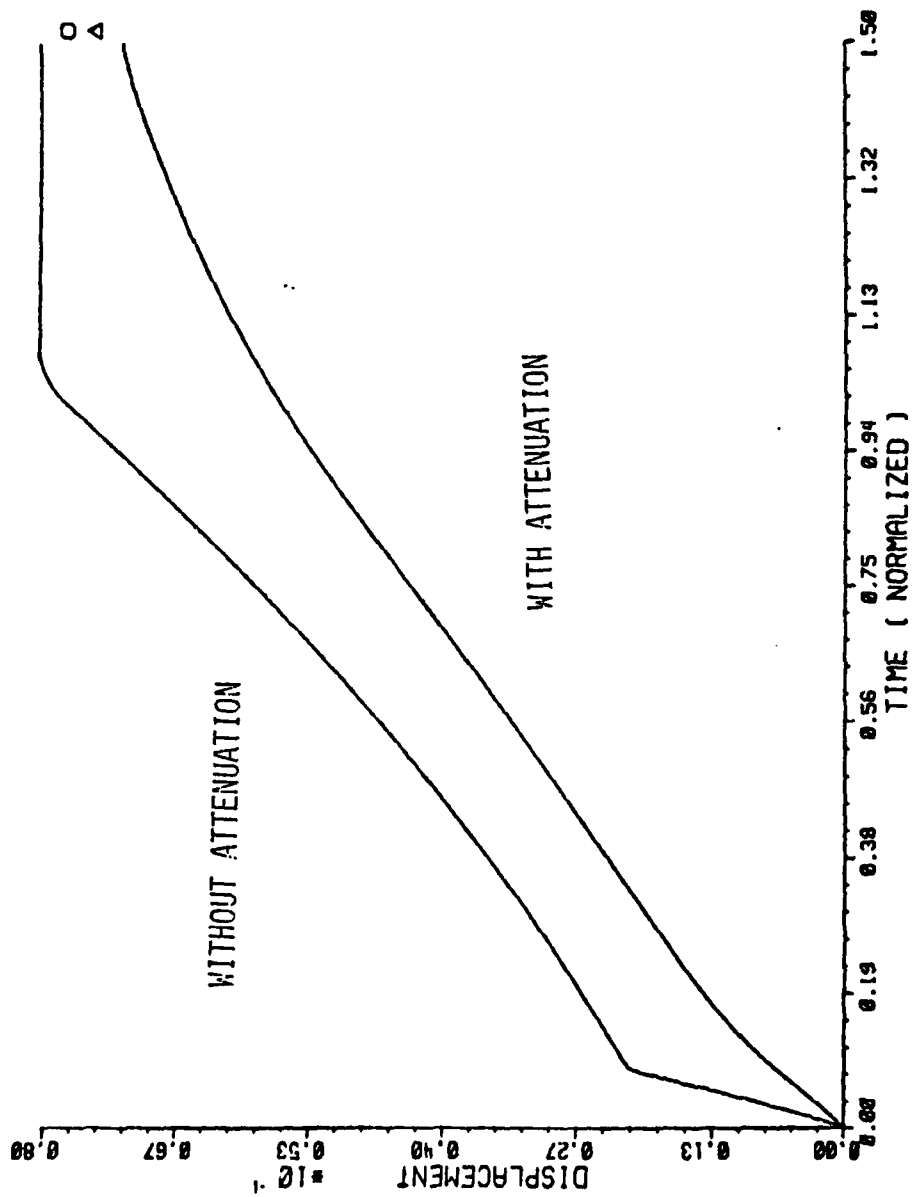


FIGURE 9. PMMA (R= 25.0 CM.)

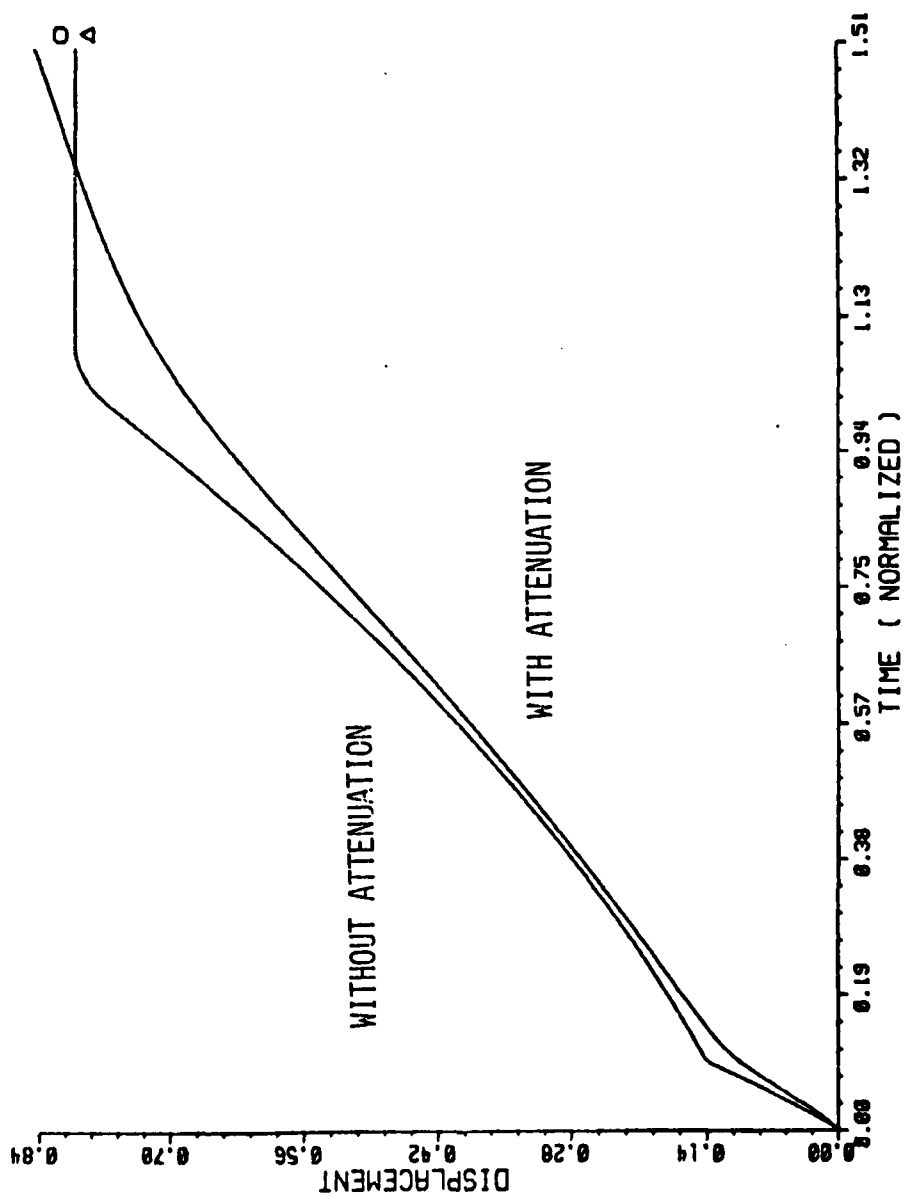


FIGURE 10. PE ( $R=2.5$  CM.)

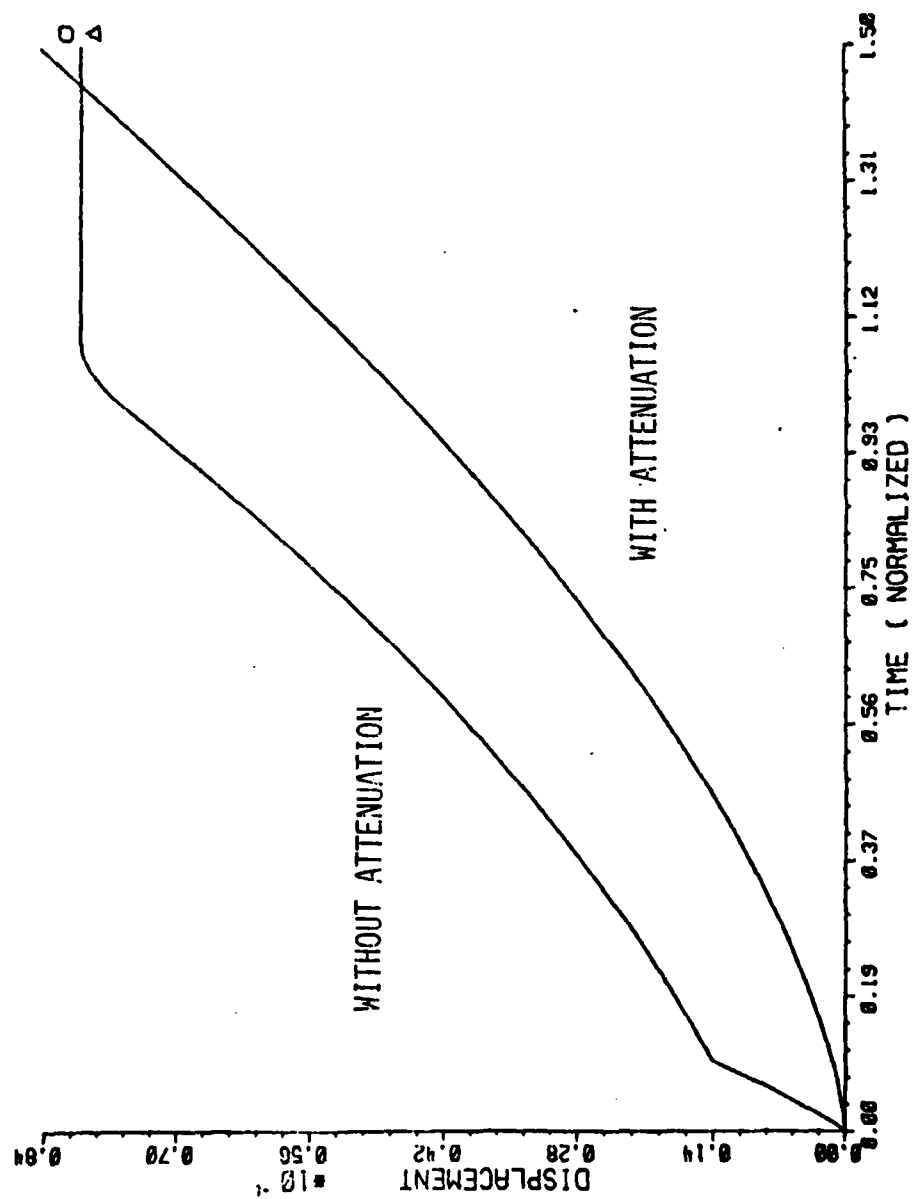


FIGURE 11. PE (R= 25.0 CM.)

source (unit amplitude, 300 NS. rise time) and propagation distances considered previously (2.5 cm. and 25.0 cm.). The differences between the specimen response for the different materials are rather significant. If one compares the response 2.5 cm. from the source in AL and PMMA for example (Figures 6 and 8) the increased amplitude reduction in PMMA readily apparent. At 25 cm. (Figures 7 and 9) the differences are even more striking. At this distance in PMMA many of the features of the waveform have been obscured making it extremely difficult to define a signal rise time. For PE where the attenuation is greater than that of PMMA the trend continues. Of particular interest is the extremely high rate of decay for shear components relative to the longitudinal in PE. This produces a situation where the shear contribution to the signal ( $U_z^s$ ) is relatively small and the displacement continues to increase well after the shear arrival. This is in sharp contrast to the situation found for aluminum, for example, as can be seen by comparing Figures 6 and 10.

To put these differences in better perspective an illustration was prepared showing the response of the three materials (at 1.0 cm.) to the same input. In Figure 12A the response is shown on a normalized basis where  $T = 0.0$  corresponds to the longitudinal arrival and  $T = 1.0$  to the shear. In Figure 12B  $T = 0.0$  is still the initial arrival, but the time scale is now real time. In the normalized presentation it appears that the initial slope of the waveform is slower for AL than for the two polymers. This is due to the difference in the ratio  $T_S/T_L$  for the different materials. Also one note the difference in signal amplitude at the "knee" of the curve due to the differences in the ratio of the shear to longitudinal wave speeds ( $C/A$ ). When the time axis is no longer normalized, this changes. For this time scale the PE signal exhibits a slower rise than the other two, principally due to the

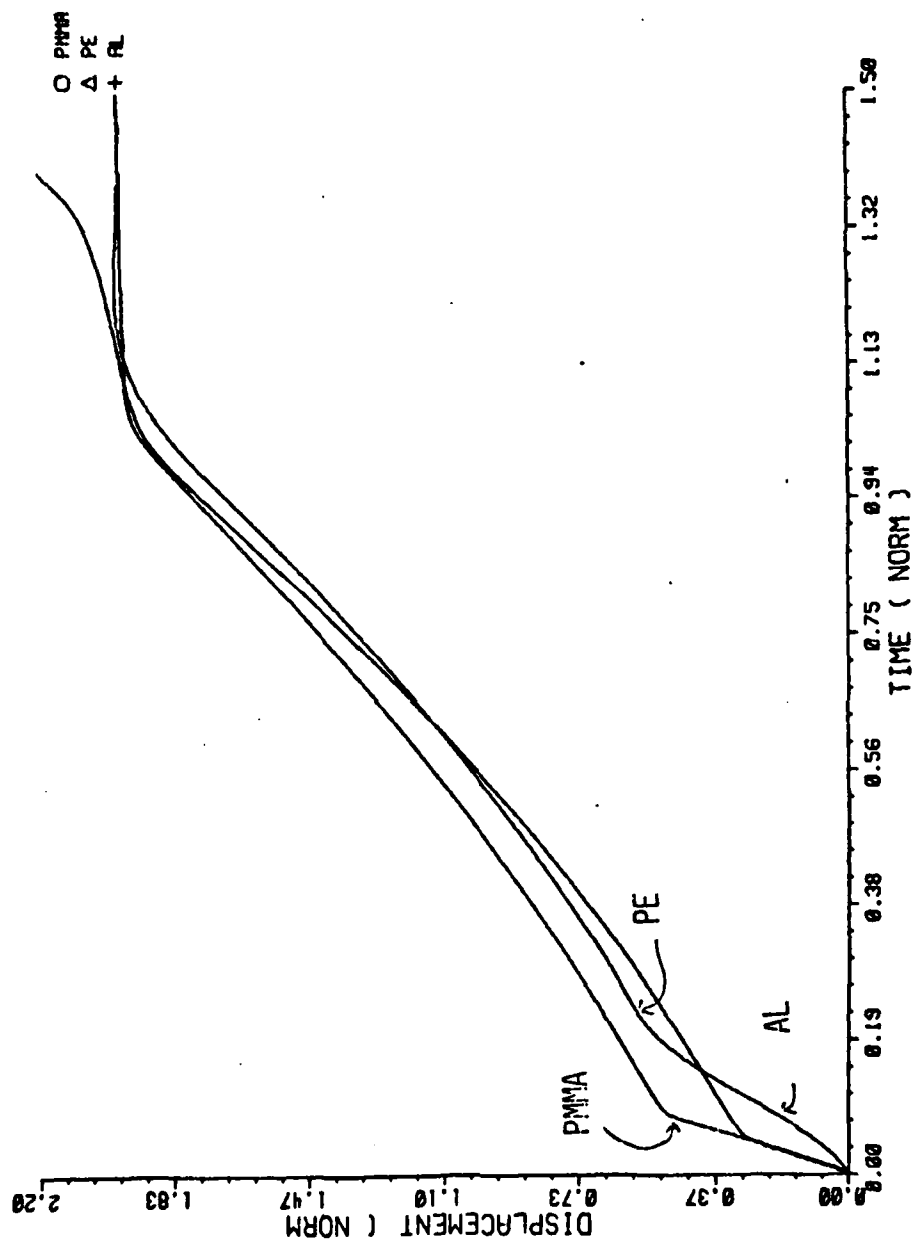


FIGURE 12A. AL, PMMA AND PE (R= 1.0 CM.) TIME AXIS NORMALIZED

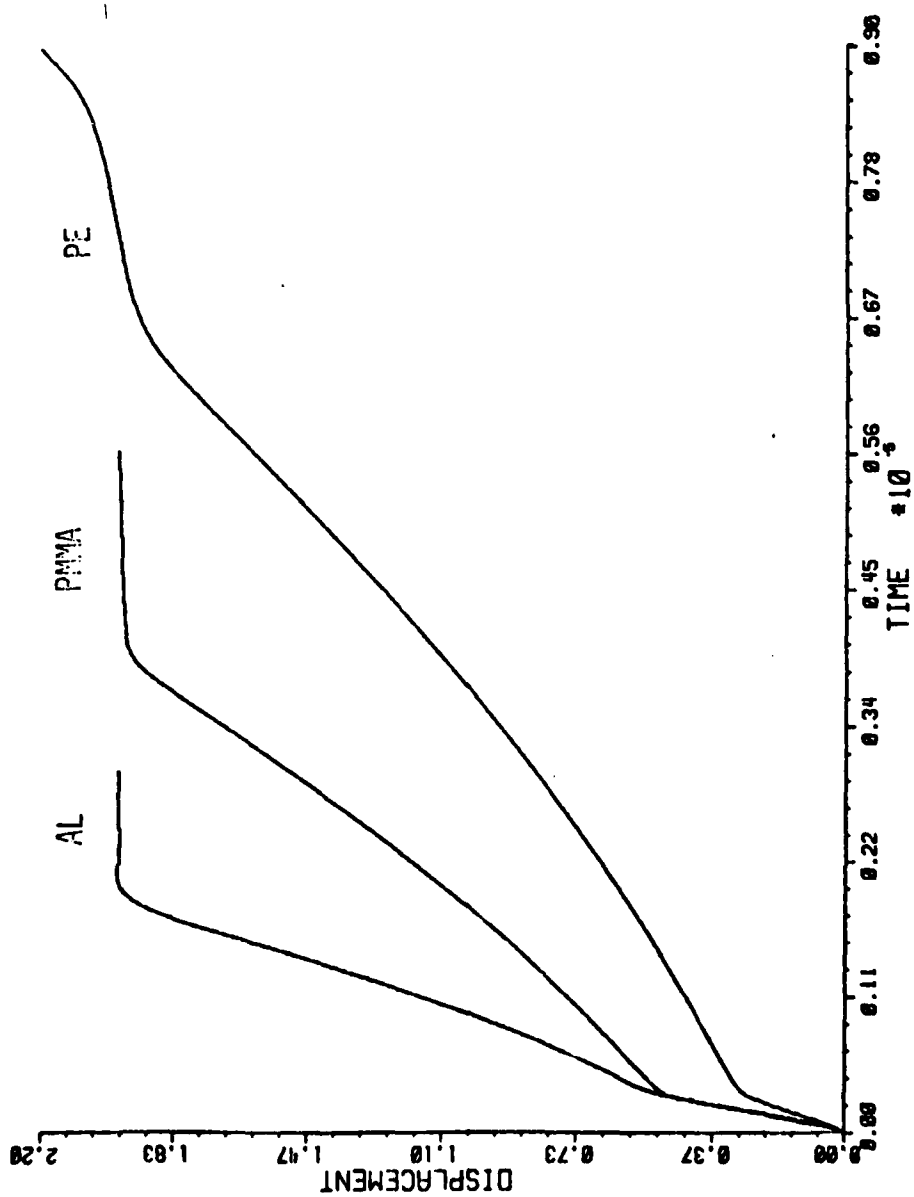


FIGURE 12B. AL, PMMA AND PE (R= 1.0 CM.) REAL TIME

slower wave speeds in this material. Figure 13 depicts the effect of changes in the source rise time on the observed displacement. Figure 14 shows the source rise time on the observed displacement. Figure 14 shows the effect of increasing propagation distance on the overall signal.

Efforts were also made to study the behavior of simulated AE signals for increasing propagation distance. Comparison of actual and predicted waveforms are presented in Figure 15 for the 0.300 cm. thick sample. Good agreement, at least qualitatively, was observed. Similar results were found for the other sample sizes investigated. It should be noted that there was a large amount of variability in signal amplitude for apparently similar situations, so it is difficult to conclusively establish quantitative connections between the model and the experiment. It is felt that the lack of signal reproducibility is attributable to using a hand held method to fracture the pencil lead rather than a fixed test jig.

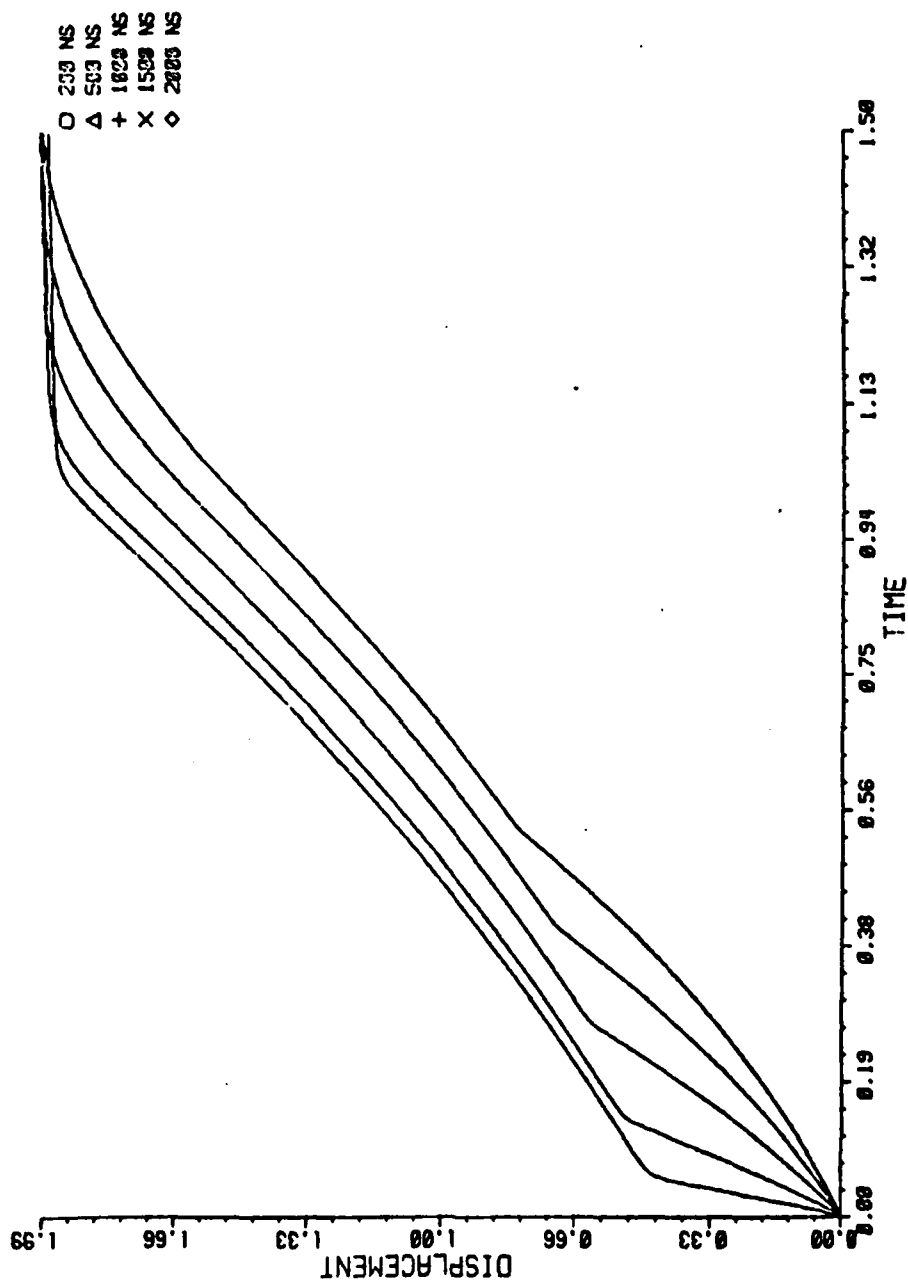


FIGURE 13. EFFECT OF SOURCE RISE TIME ON AE SIGNALS (MATERIAL PMMA)



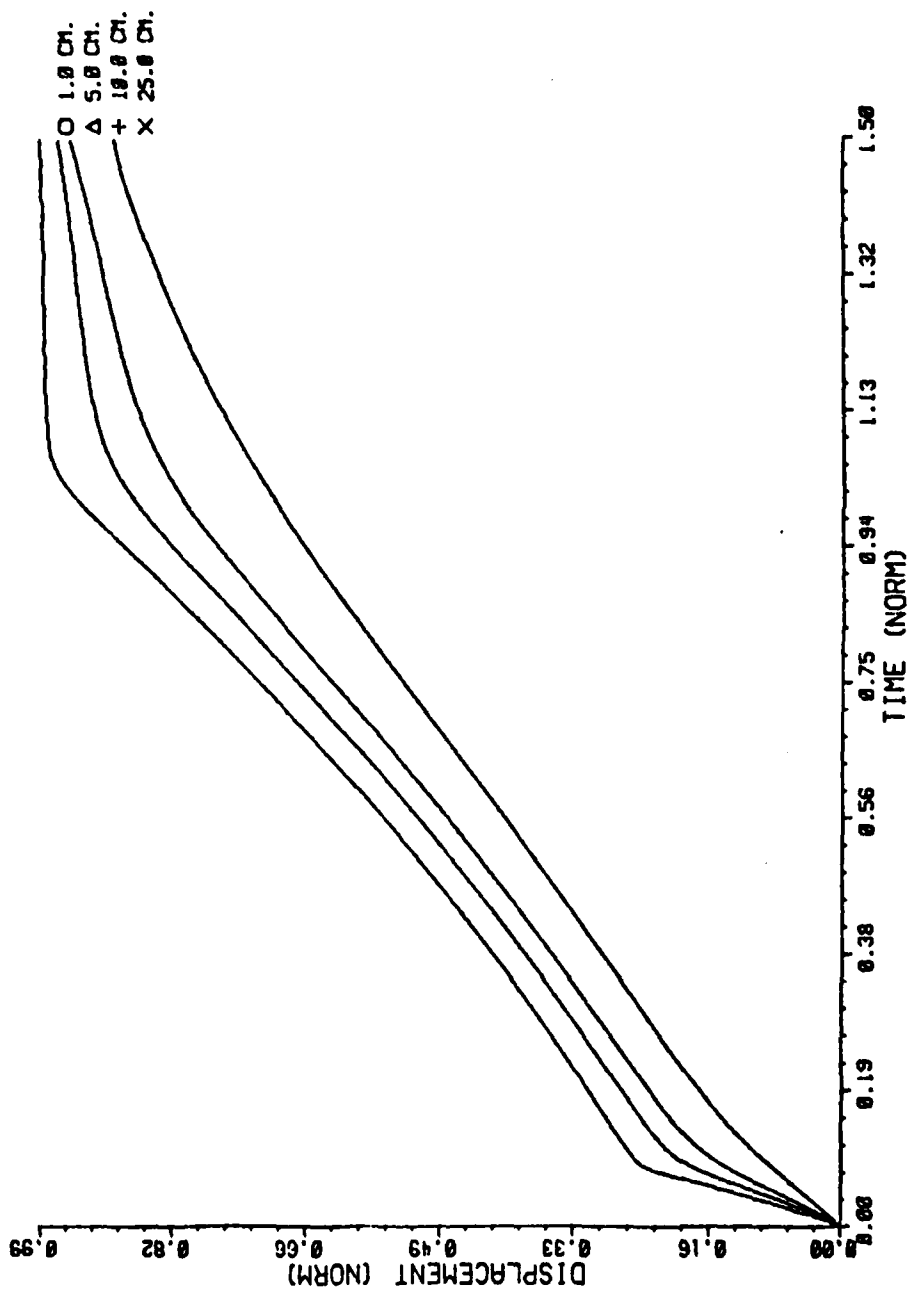


FIGURE 14. EFFECT OF PROPAGATION DISTANCE ON AE SIGNALS (MATERIAL PMMA)

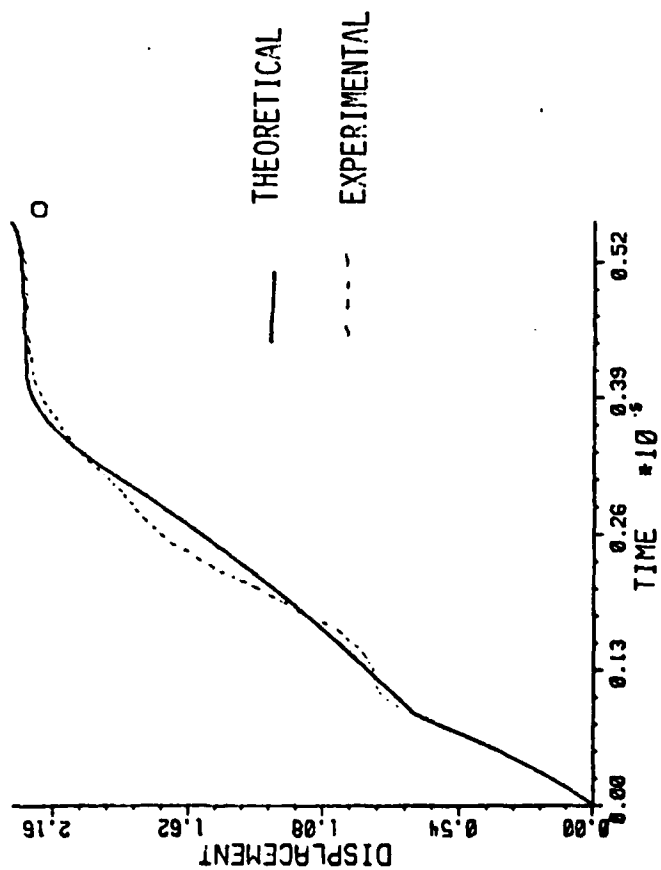


FIGURE 15. AE WAVEFORM THEORY VS. EXPERIMENT

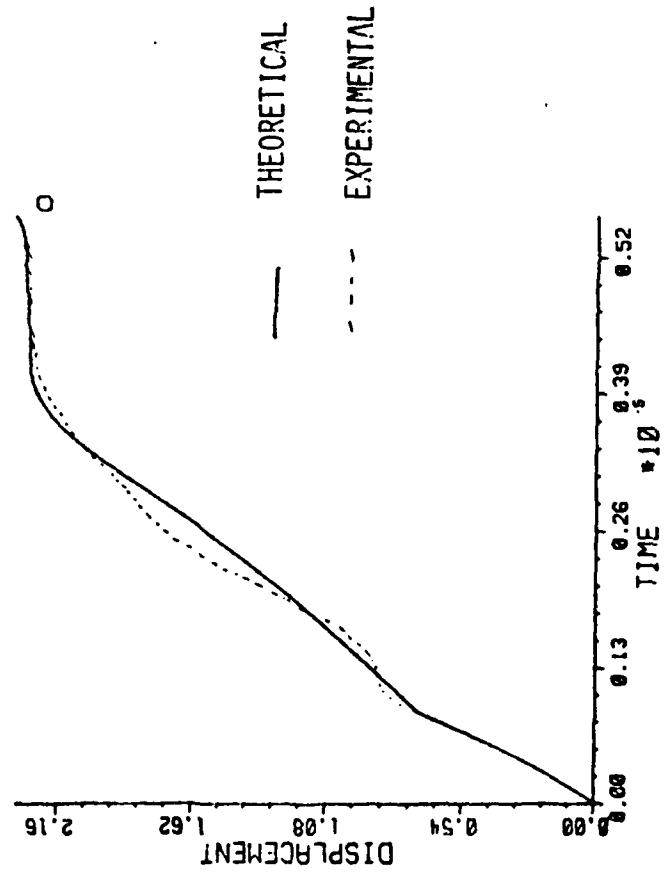


FIGURE 15. AE WAVEFORM THEORY VS. EXPERIMENT

### Conclusions

...A model has been developed to incorporate attenuation effects in predicting the response to an idealized AE source.

...The response of several different materials (AL, PMMA, PE) has been studied.

...Attenuation has been shown to produce significant deviations from earlier models in both signal rise time and amplitude.

...Experimental tests show qualitative agreement with the predictions of the theory but more testing is needed to establish the complete validity of the theory.

### Recommendations for Future Work

In this program the groundwork has been established for placing acoustic emission in composites on a sound fundamental basis. The principal theoretical advance thus far has been the introduction of material attenuation into the existing model for AE generation. This factor is of great significance in accurately determining signal characteristics (such as rise time or signal amplitude) which might permit source discrimination. Comparison of theoretical and experimental results, while qualitatively in agreement, was hampered by a lack of reproducibility in the signals and the inability to precisely measure either the force required to fracture the pencil lead or the time required for fracture to take place. It is recommended that a new system be designed to address these shortcomings. While somewhat more cumbersome, experiments with the fracture of glass capillary tubes using an indenter have proven to be more reproducible than similar experiments with lead breaks as the simulated AE source<sup>(12)</sup>. Moreover, by instrumenting the indenter with a strain gage, its force-time history can be accurately determined to permit quantitative comparison with the theoretical predictions.

## References

1. Lord, A., "Acoustic Emission-An Update," Physical Acoustics, Vol. 11 (1979).
2. Ceranoglu, A. and Y. H. Pao, "Propagation of Elastic Pulses and Acoustic Emission in a Plate," Parts I, II and III, J. Appl. Mech., 48 (1981) pp. 125-147.
3. Hsu, N., J. Simmons and S. Hardy, "An Approach to Acoustic Emission Signal Analysis-Theory and Experiment," Mat. Eval. 35 (1977) pp. 100-106.
4. Peters, R., M. Breazeale and V. Pare, "Variable Gap Detector for the Measurement of Ultrasonic Displacement Amplitudes in Solids," Rev. Sci. Instr., 39 (1968) pp. 1505-1506.
5. Breckenridge, F., M. Greenspan and C. Tscheigg, "Acoustic Emission-Some Applications of Lamb's Problem," J. Acoust. Soc. Am., 57 (1975) pp. 626-651.
6. Scruby, C. and H. Wadley, "A Calibrated Capacitance Transducer for the Detection of Acoustic Emission," J. Phys. D.: Appl. Phys., 11 (1978) pp. 1487-1494.
7. Maxfield, B. and R. Cochran, "Electromagnetic Detection of Acoustic Emission from Martensitic Transformation," Mat. Eval., 31 (1973) pp. 17-24.
8. Palmer, C. and R. Green, "Materials Evaluation by Optical Detection of Acoustic Emission Signals," Mat. Eval., 35 (1977) pp. 107-112.
9. Kline, R., R. Green and C. Palmer, "Acoustic Emission Waveforms from Cracking Steel," J. Appl. Phys., 52 (1981) pp. 141-146.
10. Scruby, C., J. Collingwood and H. Wadley, "A New Technique for the Measurement of Acoustic Emission Transients and Their Relationship to Crack Propagation," J. Phys. D.: Appl. Phys., 11 (1978) pp. 2359-2369.

11. H. Wadley and C. Scruby, "A Study of Deformation and Failure Processes in a Low Alloy Steel," *Acta Met.*, 27 (1979) pp. 613-625.
12. Michaels, J., T. Michaels and W. Sachse, "Applications of Signal Deconvolution to Acoustic Emission Signal Analysis," *Mat. Eval.*, 39 (1981) pp. 1032-1036.
13. Proctor, T., "An Improved Piezoelectric Acoustic Emission Transducer," *J. Acoust. Soc. Am.*, 71 (1982) pp. 1163-1168.
14. Pao, Y. H. and R. Gajewski, "The Generalized Ray Theory and Transient Elastic Wave Analysis in Layered Media," *Physical Acoustics*, Vol. 13 (1977).
15. Pao, Y. H., R. Gajewski and A. Ceranoglu, "Acoustic Emission and Transient Waves in an Elastic Plate," *J. Acoust. Soc. Am.*, 65 (1979) pp. 96-105.
16. Shibata, M., "A Theoretical Evaluation of Acoustic Emission Signals - The Rise-Time Effect of Dynamic Forces," *Mat. Eval.*, 42 (1984) pp. 107-116.
17. Papadakis, E., "Ultrasonic Attenuation Caused by Scattering in Polycrystalline Media," *Physical Acoustics*, Vol. 4B (1968).
18. Hartmann, B. and J. Jarzynski, "Ultrasonic Hysteresis Absorption in Polymers," *J. Appl. Phys.*, 43 (1972) pp. 4304-4372.

# Appendix 1: AE Surface Displacements

The surface displacement terms (in the absence of attenuation)  $U_z^L$  and  $U_z^S$ , are given by

$$U_z^L = \begin{cases} 0 & t < t_L \\ C^2 \left[ \frac{2t}{\tau} (C_1 + C_2 (t^2 - t_L^2)) - \frac{2}{\tau} (C_1 t_2 + \frac{2}{3} (t^3 - t_L^3)) \right] & t_L < t < t_L + \tau \\ C^2 \left[ \frac{2t}{\tau} C_2 (t^2 - t_L^2) - \frac{2}{\tau} \left( \frac{2}{3} C_2 (t^3 - \bar{t}_L^3) \right) + 2 C_1 + C_2 ((t^*)^2 - t_L^2) \right] & t_L + \tau < t \end{cases}$$

$$U_z^S = \begin{cases} 0 & t < t_S \\ C^2 \left[ \frac{2t}{\tau} C_2 (t^2 - t^{*2}) - \frac{2}{\tau} \left( \frac{2}{3} C_2 (t^3 - t^{*3}) \right) \right] & t_S < t < t_S + \tau \\ C^2 \left[ \frac{2t}{\tau} C_2 (t^2 - t^{*2}) - \frac{2}{\tau} \left( \frac{2}{3} C_2 (t^3 - t^{*3}) \right) + 2 (C_1 + C_2 (t^{*2} - t_S^2)) \right] & t_S + \tau < t \end{cases}$$

where

$C$  = Shear Wave Speed

$C_1 = 1/A^2$

$C_2 = 1/R^2$

$\bar{t}_S = t - t_S$

$\bar{t}_L = t - t_L$

$t^* = t - \tau$



# APPENDIX 2.

```

DIMENSION DEXY(1024),XVSC(1024)
VIRTUAL AMOS(1024),ARG(1024),FR(1024),Z1(1024)
VIRTUAL TEM(1024),TEM1(1024)
INTEGER SGU
000 LONGITUDINAL DISP (TEM)
000 SHEAR DISP (TEM1)
000 LONG-SHEAR DISP (Z1)
CALL ASSIGN (4,'TT1: ')
WRITE(5,100)
100 FORMAT(1X,'LONGITUDINAL WAVE SPEED(E10.4)= ')
READ(5,101)A
101 FORMAT(E10.4)
WRITE(5,105)
105 FORMAT(1X,'S-SHEAR WAVE SPEED(E10.4)= ')
READ(5,101)C
WRITE(5,110)
110 FORMAT(1X,'SOURCE RISE TIME(E10.4)= ')
READ(5,101)TAU
WRITE(5,115)
115 FORMAT(1X,'SPECIMEN THICKNESS(E10.4)= ')
READ(5,101)R
R1=R
N=1024
DT=R1*1.0E-8
TMAX=DT*FLOAT(N)
PI=3.1415927
PMAX=(FLOAT(N)*2*PI*R1)/TMAX
TL=R/A
TS=R/C
P=A/C
P1=P**2
C1=1.0/(A**2)
C2=1.0/(R**2)
000 CALCULATION OF SURFACE DISPLACEMENT COMPONENTS
000 AND STORE THEM IN ARRAYS TEM AND TEM1
DO 200 I=1,N
T=0.0
T=FLOAT(I)*DT
IF (T-TL) 201,202,202
201 A1=0.
A2=0.
A3=0.
B1=0.
B2=0.
B3=0.
GOTO 200
202 IF (T-TL-TAU) 203,203,204
203 A1=C1+C2*((T**2)-(TL**2))
A2=C1*TL+(2.0/3.0)*C2*((T**3)-(TL**3))
A3=0.0
B1=0.0
B2=0.0
B3=0.0
GOTO 200
204 IF (T-TS) 205,205,206
205 TLO=T-TAU
A1=C2*((T**2)-(TLO**2))
A2=(2.0/3.0)*C2*((T**3)-(TLO**3))
A3=C1+C2*((TLO**2)-(TL**2))

```

```

      B3=0.0
      B3=2.0
      GO TO 600
206  IF (T-TS-TAU) 207, 208, 208
207  TLO=T-TAU
      A1=C2*(T**2)-(TLO**2)
      A2=(2.0/3.0)*C2*(T**3)-(TLO**3)
      A3=C1+C2*(T-TAU)**2-TL**2
      B1=C2*(T**2-TS**2)
      B2=(2.0/3.0)*C2*(T**3-TS**3)
      B3=0.0
      GO TO 320
208  TLO=T-TAU
      A1=C2*(T**2-TLO**2)
      B1=C2*(T**2-TLO**2)
      A2=(2.0/3.0)*C2*(T**3-TLO**3)
      B2=(2.0/3.0)*C2*(T**3-TLO**3)
      A3=C1+C2*(TLO**2-TL**2)
      B3=C2*(TLO**2-TS**2)

320  TEM(I)=(C**2)*((2.0*T/TAU)*A1-(2.0/TAU)*A2+(2.0*A3))
      TEM1(I)=(C**2)*((2.0*T/TAU)*B1-(2.0/TAU)*B2+(2.0*B3))
      TEM(I)=TEM(I)/R1
      TEM1(I)=TEM1(I)/R1
      B2=TEM(N)-TEM1(N)
      XVEC(I)=FLOAT(I-1)*DT
200  CONTINUE
      DO 800 IS=1,3
      IF (IS-2) 810, 820, 830
810  DO 710 NN=1,N
      DUMY(NN)=TEM(NN)
710  CONTINUE
      GO TO 750
846  DO 847 NC=1,N
      TEM(NC)=DUMY(NC)
847  CONTINUE
      GO TO 800

820  DO 730 NX=1,N
      DUMY(NX)=TEM1(NX)
730  CONTINUE
      GO TO 750
849  DO 850 ND=1,N
      TEM1(ND)=DUMY(ND)
850  CONTINUE
      GO TO 800

830  DO 720 NI=1,N
      Z1(NI)=TEM(NI)-TEM1(NI)
      DUMY(NI)=Z1(NI)
720  CONTINUE
      CALL PLOTIN(TMAX,B2)
      CALL PLOTXY(XVEC,DUMY,N)
      PAUSE TYPE RES TTS TO CONTINUE
800  CONTINUE

      GO TO 860
750  B4=DUMY(N)
      CALL PLOTIN(TMAX,B4)
      CALL PLOTXY(XVEC,DUMY,N)
      PAUSE TYPE RES TTS TO CONTINUE
      DO 735 IA=1,N
      XVID(IA)=0.0
735  CONTINUE
      NCON=

```

```

      CALL FFTS(DUMY, XVEC, N, NCP, SGN)
      DO 735 II=1, N
        DUMY(II) = DT*DUMY(II)
        XVEC(II) = DT*XVEC(II)
735  CONTINUE
      N3=N/2+1
      DO 735 II=1, N3
        FR(II) = (FLOAT(II-1)*2*PI*R)/TMAX
738  CONTINUE
      IF (IS.EQ.2) GOTO 920
      WRITE(5,901)
901  FORMAT(1X, 'ATTENUATION CONST K1(LONG), E10.4=')
      READ(5,101)AL1
      WRITE(5,903)
903  FORMAT(1X, 'ATTENUATION CONST K2(LONG), E10.4=')
      READ(5,101)AL2
      DO 910 II=1, N3
        DUMY(II) = DUMY(II)*EXP(-R1*(AL1*FR(II)+AL2*(FR(II)**4)))
        XVEC(II) = XVEC(II)*EXP(-R1*(AL1*FR(II)+AL2*(FR(II)**4)))
        DUMY(N-II+2) = DUMY(II)
        XVEC(N-II+2) = -XVEC(II)
910  CONTINUE
      GOTO 931
920  WRITE(5,921)
921  FORMAT(1X, 'ATTENUATION CONST K1(SHEAR), E10.4=')
      READ(5,101)AS1
      WRITE(5,923)
923  FORMAT(1X, 'ATTENUATION CONST K2(SHEAR), E10.4=')
      READ(5,101)AS2
      DO 930 IO=1, N3
        DUMY(IO) = DUMY(IO)*EXP(-R1*(AS1*FR(IO)+AS2*(FR(IO)**4)))
        XVEC(IO) = XVEC(IO)*EXP(-R1*(AS1*FR(IO)+AS2*(FR(IO)**4)))
        DUMY(N-IO+2) = DUMY(IO)
        XVEC(N-IO+2) = -XVEC(IO)
930  CONTINUE
931  NCP=1
      SGN=-1
      CALL FFTS(DUMY, XVEC, N, NCP, SGN)
      DO 745 ID=1, N
        DUMY(ID) = DUMY(ID)/TMAX
        IF (DUMY(ID).LT.0.0) DUMY(ID) = 0.0
        XVEC(ID) = FLOAT(ID-1)*DT
745  CONTINUE
      CALL PLOTIN(TMAX, B4)
      CALL PLOTXY(XVEC, DUMY, N)
      PAUSE 'TYPE RES ITS TO CONTINUE'
      IF (IS.EQ.1) GOTO 846
      GOTO 849
860  DO 700 MY=1, N
      WRITE(1,600)TEM(MM), TEM1(MM), Z1(MM), XVEC(MM), MM
700  CONTINUE
600  FORMAT(1X, 4E14.4, 1X, I4)
      STOP
      END
      SUBROUTINE PLOTIN(TAM, R5)
      CALL CLEAR
      CALL SCALE(-TAM*0.1, 1.05*TAM, -0.1*RS, 1.2*RS)
      CALL XAXIS(0., 0., TAM, TAM/10.0, 1.)
      CALL NUMHX(0., 0., TAM)
      CALL HLABEL(TAM/2.0, -0.05*RS, 'TIME (SEC)', 10)
      CALL YAXIS(0., 0., RS*1.1, RS/10.0, 1.)
      CALL NUMHY(0., 0., 1.1*RS)
      CALL VLABEL(-0.05*TAM, 0.5*RS, 'DISPLACEMENT', 10)
      RETURN
      END

```

## PART II

### Acoustic Emission in a Single Fiber Reinforced Composite

#### BACKGROUND

The elastic strain energy stored when a graphite fiber reinforced plastic (GFRP) material is under load may be released in several ways. While some of the strain energy released produces transient elastic stress waves which propagate through the material, additional energy is in the form of heat, surface strain energy and plastic strain energy. The three fundamental mechanisms which generate these elastic stress waves in a GFRP material are 1) fiber fracture, 2) matrix cracking and 3) interfacial debonding. As these stress waves reach the surface of the material (from its buried source), a transducer can be used to detect the particle displacement of the stress waves. These detected stress waves are known as acoustic emissions.[1-2]

It is believed that the acoustic emission wave form is characteristic of its emission source. However, the detected signal is also a function of propagational effects such as the material type, material size and geometry. In addition, the waveform may be modified by the detection system (transducer, coupling, electronic apparatus)[1-3]. Though it has just been stated that the acoustic emission signal is a function of factors other than the source of the signal, certain geometrical requirements as well as a broad bandwidth transducer (sufficiently larger than the source bandwidth) should permit part of the source event to be recorded [2]. Geometrical requirements suitable to minimize internal reflections include locating the transducer as close to the source as possible and a geometry such that surface reflections are directed away from the transducer.

A great deal of research has been spent trying to correlate acoustic emission features with specific fracture mechanisms in solids. Although most

of the work in this area has been done with metals, more research is being devoted to composite materials in the recent decade.

Speake and Curtis [4] experimented with carbon fiber reinforced plastic specimens. Their results showed the AE signal to be strongly dependent upon specimen size, geometry and jig mounting. They did however find some characteristic frequencies of acoustic emissions from different fiber types.

Henneke and Russell [5] experimented with single or three ply graphite-epoxy, 1 inch coupons. They found the emission spectra from fiber breakage to be predominately at higher frequencies, while matrix cracking was at lower frequencies. They believe, as we do, that fiber breakage models cannot be accurately evaluated in real composite laminates until the waveform for a single fiber breakage is characterized.

Mehan and Mullin [6] studied individual fiber failure of boron filaments in an epoxy matrix specimen. They found each failure mechanism in a given composite to have different acoustic emission signatures. The total frequency characteristic of the signal was analyzed. This method has the drawback that the long-term signal is a function of the specimen size, geometry and normal mode of specimen vibration, as well as the failure mechanism.

Narisawa and Oba [7] investigated composite specimens containing a single filament carbon fiber in an epoxy matrix. The AE over the life of the specimen was studied, whereby the life of the specimen included multiple fiber breaks in the fiber. They found the AE burst to have one to one correspondence with the fiber fracture. Signal analysis was not carried out in this study.

The basic test specimen in this experiment will consist of a single filament graphite fiber in an epoxy matrix. Unlike previous work done with composite laminates, the use of a single filament fiber with suitable matrix

geometry will enable the location of fiber failure to be kept constant. Although other work has been done with single fiber specimens, this experiment differs in that the matrix geometry and transducer location will allow the leading edge of the stress wave to be deflected before multiple reflections reach the surface and further distort the signal.

By identifying the characteristic signature of an acoustic emission event in a single filament GFRP specimen, it may someday be possible to identify specific critical events in proof tests of complex composite structures such as those used in the aerospace industry. After identification of failure mechanisms, the engineer can then redesign future structures to strengthen a particular component of the composite structure which had failed in proof tests. Furthermore, in the case of large, very expensive composite structures, the location of the critical event can be determined and repaired to preserve the entire structure. One can easily see the impact this type of study may someday have on the composite materials industry.

It is the intent of this experimental research to identify characteristic acoustic emission waveforms as they relate to the emission source. Fiber breakage (in a graphite fiber reinforced epoxy matrix) will be the principle failure mechanisms studied. Furthermore, through suitable specimen size and geometry, the leading edge of the waveform will be examined before the wave is distorted by internal specimen reflections. From this, the rise-time and amplitude will be correlated to the characteristics of the failure mechanism.

#### Test System

The test specimen is loaded in tension to failure. The tensile loading jig (see Fig. 1, for tensile loading jig) is loaded manually at a near constant rate. The transducer is placed vertically above the emission source

at epicenter. The position of the transducer relative to the tensile loading jig can be seen in Figure 2. A 2.48 lb lead weight is centered on top of the transducer to insure contact is maintained during specimen fracture.

During deformation of the test specimen, acoustic emission events are detected with a NBS/PCT acoustic emission AE transducer. This AE transducer is damped such that its bandwidth is relatively flat over the bandwidth of interest (0.01-10MHz). As can be seen in Figure 3, the transducer's piezoelectric crystal is conical. With only the tip of the cone contacting the surface, the signal detected more closely approximates a point source receiver than conventional transducer. Furthermore, the backing is large in the radial and axial direction and reduces the effect of transducer resonance. Finally, the small contact area reduces interference due to the aperture effect from stress waves approaching the transducer at angles not normal to the transducer [8].

The acoustic emission detection system shown in Figure 4 uses a transducer couples to a high-gain differential amplifier and filter.\* The band-pass filter was set at low pass (DC) high frequency (1MHz). The counter threshold voltage is 20mV. The signal travels to a transient signal digitizer. The digitizer\*\* is set such that it measures at  $\pm 10$  volts full scale, 50 nsec. time per point through an external source. The trigger was set such that low level emissions would not trigger, while larger emissions such as those by matrix cracking and fiber breakage would trigger. The waveform was stored on diskette from which plots were generated at a later time.

\*Tecktroni, Inc., Type 1A7A High-Gain Differential Amplifier.

\*\*Nicolet Instrument Corp., Model 204-A.

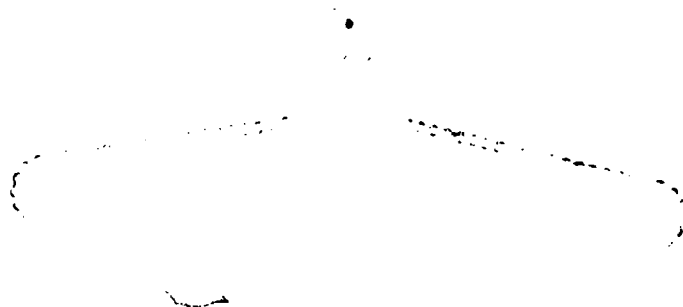


Figure 1. Tensile loading jig.

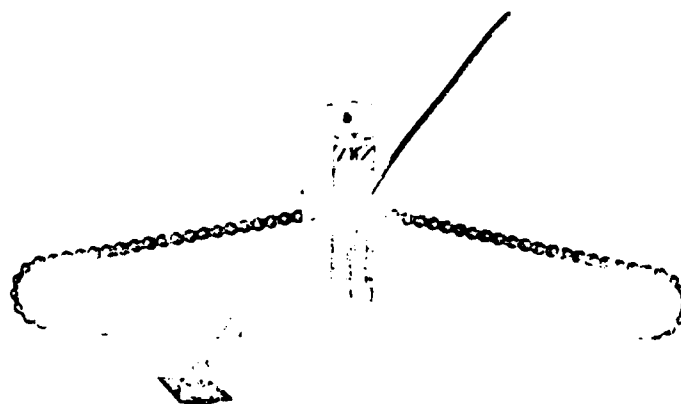


Figure 2. Jig with transducer mounted.



Figure 3. NBS/PCT Acoustic Emission Transducer.

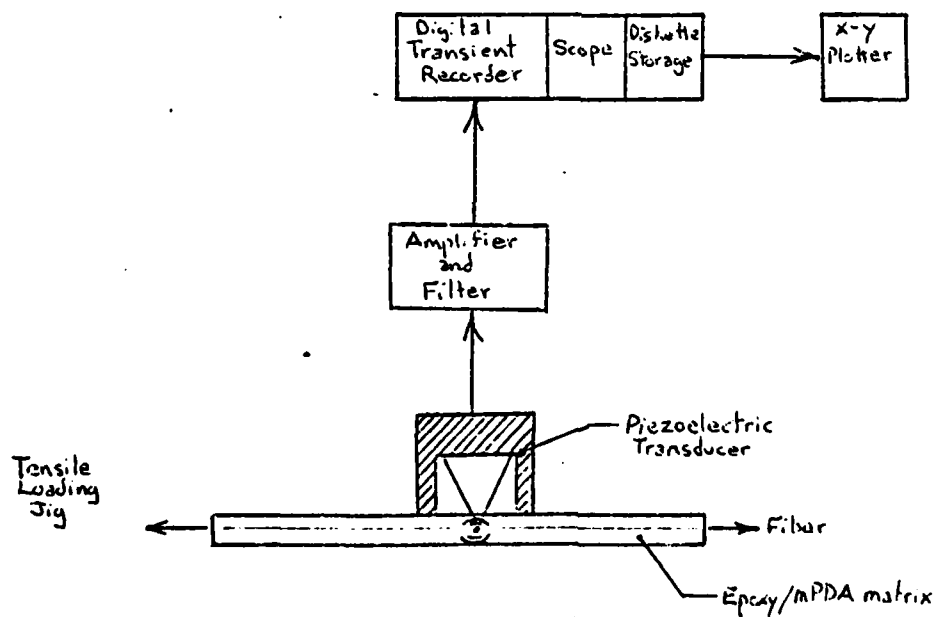


Figure 4. Acoustic Emission detection system.

Three test specimens were used during the experimental survey. All specimens had the same epoxy/MPDA matrix. In addition, two of the three specimens tested included a single filament graphite fiber buried longitudinally in the matrix. A saw cut using a diamond wheel was placed at the mid-section of the specimen to insure matrix and/or fiber failure at a localized source. This enabled us to place the transducer at epicenter with the buried source (see Fig. 5 for specimen geometry).

The matrix used during this survey was a two-part epoxy system. EPON 828 (diglycidal ether of bisphenol-A) was used as the epoxy resin, while MPDA (meta phenylene diamine, 99+%) was added at 14.5 phr as a curing agent. The epoxy/MPDA matrix was vacuum degassed for seven minutes to reduce the presence of bubbles in the matrix. Following degassing the matrix was cured in an oven

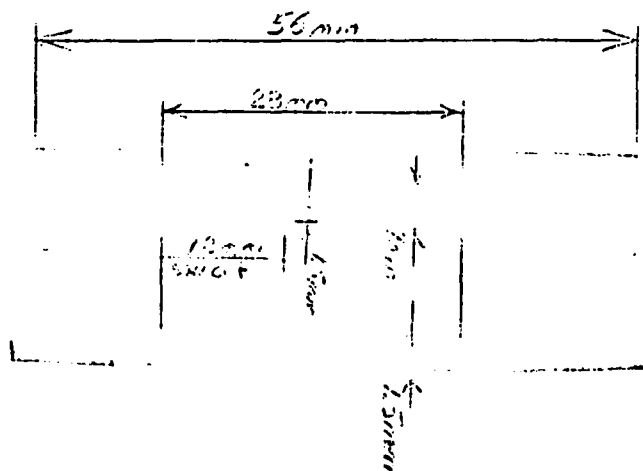


Figure 5. Single filament graphite fiber reinforced plastic test specimen.

for two hours at 75°C followed by an additional two hours at 125°C and furnace cooling. The cured matrix has a tensile modulus of approximately 400,000 psi and tensile strength of approximately 11,000 ps. [9-10].

Two of the three specimen types included a single filament graphite fiber, placed longitudinally in the epoxy/MPDA matrix. The two fibers chosen for the survey were selected because of their two distinct mean diameters and mechanical properties. One fiber (hereafter denoted AS4) has a nominal tensile modulus of  $33 \times 10^6$  psi, tensile strength of  $565 \times 10^3$  psi and a 7.6  $\mu\text{m}$  mean diameter. The other fiber (AS6) has a tensile modulus of  $34.4 \times 10^6$  psi, tensile strength of  $604 \times 10^3$  psi and a 48  $\mu\text{m}$  mean diameter. The fibers were without surface treatments [9-10].

#### Results and Discussion

The specimen used in this experiment is a single filament graphite fiber/epoxy matrix specimen with a saw cut at the specimen mid-section so that deformation can be localized. The principle failure mechanism initiates with fiber fracture. This geometry also permits the location of the fiber failure to be kept constant. By knowing the location at which the fiber will fail, it is possible to place the transducer directly above the emission source at epicenter. Signals measured at epicenter from this type of source are simplified because they do not include the surface waves.

The geometry of the specimen and placement of the transducer relative to the emission source allowed the first dilational wave peak to be detected and recorded before the wavefront was heavily distorted by internal reflections. From the data collected, it was observed that two distinctly different signatures resulted from epoxy matrix cracking and that of fiber breakage. The two waveforms can be seen in Figure 6 and 7. Notice that only the leading edge

of the pulse is characteristic of the failure mechanisms. The remainder of the pulse is dominated by the normal modes of vibration of the specimen and internal reflections of the stress wave.

Two measurements were made from the first peak of the stress wave in order to correlate the stress wave to the finite dimension of the failure mechanism. The two measurements were the amplitude and the rise time. The rise time, shown in Figure 8, is a measure of the duration of the failure event, while the amplitude is related to the magnitude of force of the event [1].

The mean rise time for the specimen with the AS4 fiber was 9.03  $\mu$ s. This was calculated from twelve test specimens, all of which had the same characteristic signature. The standard deviation was  $\pm 1.49$   $\mu$ s from the mean. The mean amplitude was likewise calculated to be 7.22 volts,  $\pm 1.72$  v standard deviation from the mean. The mean rise time for the specimen with the AS6 fiber (from a thirteen specimen lot all with the same characteristic signature) was 9.72  $\mu$ s,  $\pm 1.17$   $\mu$ s standard deviation from the mean. The mean amplitude was 8.42 volts,  $\pm 2.26$  v standard deviation from the mean. Because the AS6 fiber has a slightly higher tensile strength than the AS4 (see text for mechanical properties), it was expected that the amplitude of the AS6 signal would be higher, since the fiber diameters did not vary significantly. However, since the AS6 fiber (5.8  $\mu$ m) was slightly less than the AS4 fiber (7.6  $\mu$ m) it was expected that the AS6 would have a similar but slightly lower rise time than the AS4 fiber. This was not the case. Although this result cannot be fully explained, it could be caused by the AS6 fiber being a softer fiber than the AS4 and thus causing a slower crack growth. Furthermore, by direct observation of the waveform, it appears as though there was some ringing in the specimen.

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

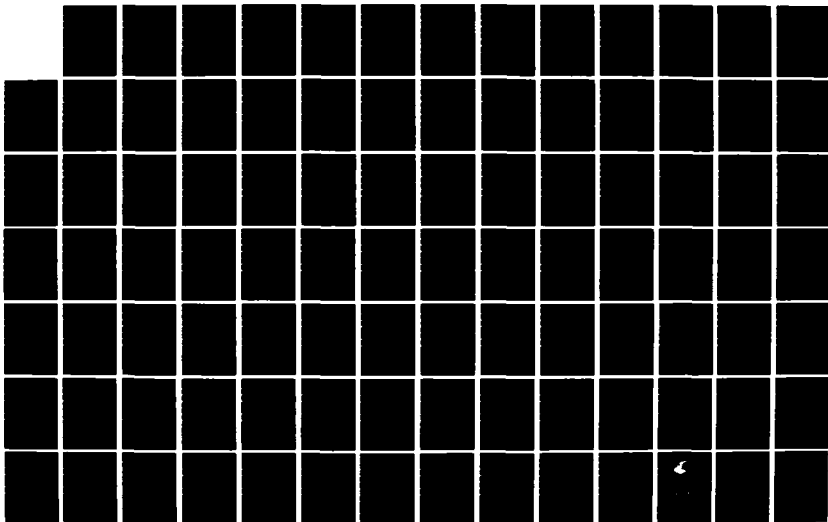
8/13

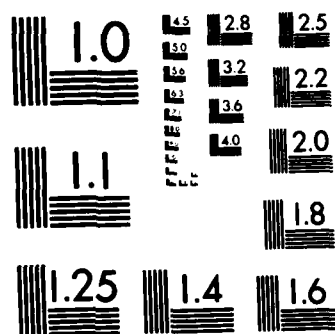
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

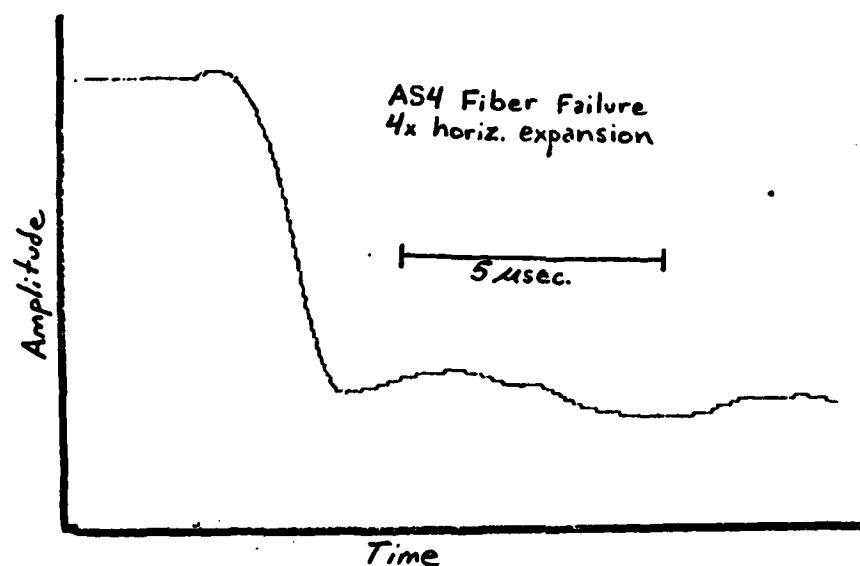


Figure 6a. Leading edge of AE waveform, fiber failure.

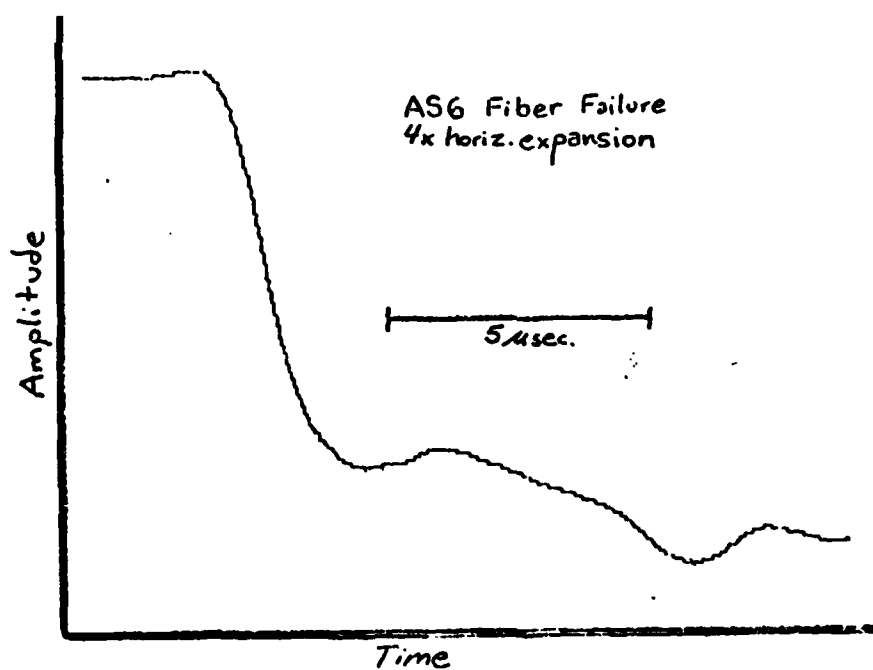


Figure 6b. Leading edge of AE waveform, fiber failure.

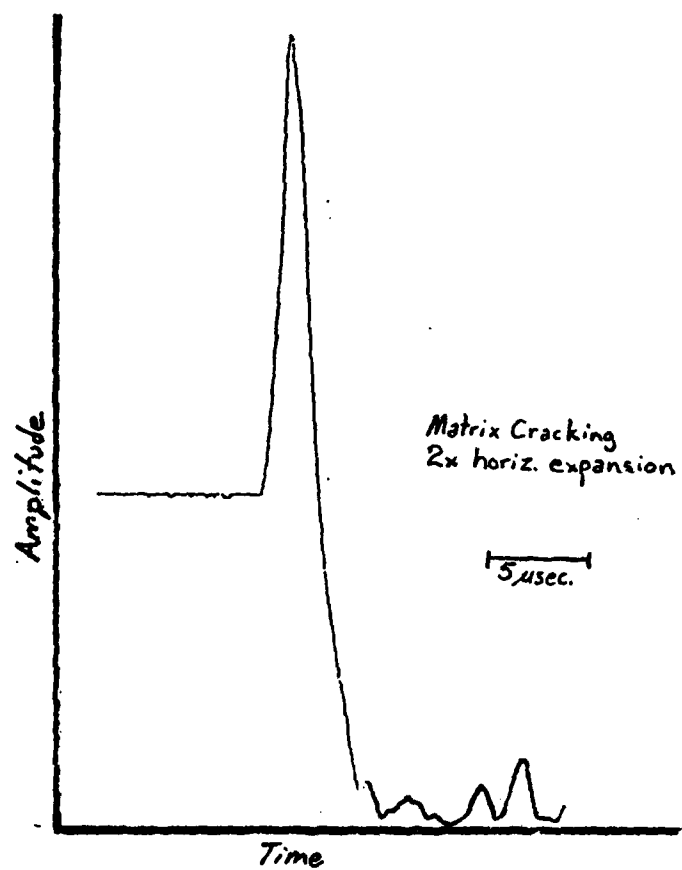


Figure 7. Leading edge of AE waveform, matrix cracking.

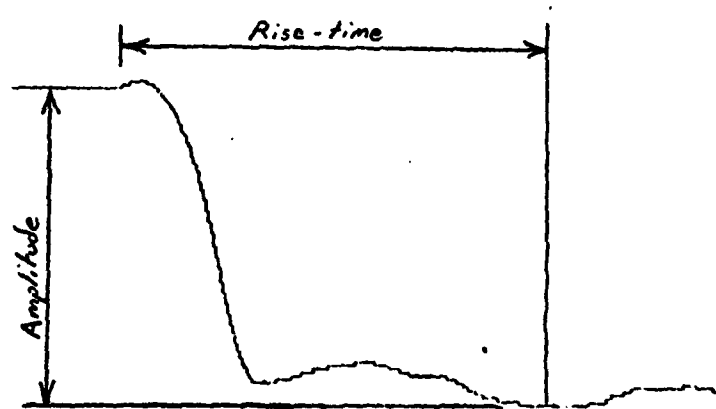


Figure 8. A typical waveform showing parameters referenced in text.



As previously mentioned, matrix cracking had a distinctly different signature than that of fiber breakage. The rise time for matrix cracking was calculated from nineteen specimens which all showed the same characteristic signature. The mean rise time was  $6.52 \mu s$ ,  $\pm 1.34 \mu s$  standard deviation. The mean amplitude was 7.6 volts,  $\pm 3.58 v$  standard deviation. Note that the standard deviation is much greater than the other results. This can be attributed to the fact that the saw cut varied substantially with each specimen and caused different loads to initiate matrix cracking.

By improving the reproducibility of the saw cut on the specimen and measuring the load at failure, the deviation from the mean could be reduced and/or better explained. Furthermore, changing the specimen size, geometry and jig mounting could further improve the reproducibility of the waveforms.

#### Recommendations

This experimental research generated two distinctly different waveform for two failure mechanisms in a graphite fiber reinforced epoxy specimen. The failure mechanisms identified are fiber breakage and matrix cracking. Only the leading edge of these waveforms should be considered characteristic of the failure modes, since the long-term waveform is more characteristic of the specimen size, geometry and jig mounting than the failure modes. The mean amplitude of the matrix cracking is less than the mean amplitude for fiber failure. The mean rise-time is faster for matrix cracking than for fiber breakage.

Because the standard deviations from the mean amplitudes were slightly higher than we would have liked, the following is recommended for further work in order to improve reproducibility of results and to assist in explaining the current results:

- 1) Modify specimen size and geometry such that the effect of specimen resonance is minimized and the arrival time of reflected stress waves is increased.
- 2) Measure the strain and load during specimen failure in order to better explain varying rise-times and amplitudes.
- 3) Improve the reproducibility of the saw cut or design a specimen such that no saw cut is necessary to localized deformation at a constant location vertically below the transducer.
- 4) Experiment with additional fibers with more different fiber diameters and tensile strengths.

Although the experiment was successful in producing waveforms recognizably different and characteristic of the emission source, more work needs to be done relative to the recommendations discussed herein.

### References

- [1] Scruby, C. B., Collingwood, J. C., and Wadley, H. N. G., "A New Technique for the Measurement of Acoustic Emission Transients and Their Relationship to Crack Propagation." *J. Phys. D* 11 (1978): 2359-2369.
- [2] Kline, R. A., Palmer, C. H., and Green Jr., R. E., "Acoustic Emission Waveforms from Cracking Steel: Experiment and Theory." *J. Appl. Phys.* 52 (1981): 141-146.
- [3] Shibata, M., *Materials Evaluation* 42 (1984) 107.
- [4] Speake, J. H., and Curtis, G. J., "Characterization of the Fracture Processes in CFRP Using Spectral Analysis of the Acoustic Emissions Arising from the Applications of Stress." *International Conference on Carbon Fibers*, London, February 1974. The Plastics Institution, Paper No. 29.
- [5] Henneke, E. G., and Russell, S. S., "Signature Analysis of Acoustic Emission from Graphite/Epoxy Composites." *National Aeronautics and Space Administration Report No. VPI-E-77-22*, September 1977.
- [6] Mehan, R. L., and Mullin, J. V., "Evaluation of Composite Failure through Fracture Signal Analysis." *J. Testing and Evaluation* 1 (1973): 215-219.
- [7] Narisawa, I. and Oba, H., "An Evaluation of Acoustic Emission from Fibre-reinforced Composites." *J. Materials Science* 19 (1984): 1777-1786.

- [8] Proctor Jr., T. M., "An Improved Piezoelectric Acoustic Emission Transducer." J. Acoust. Soc. Am. 71 (1982): 1163-1168.
- [9] Drzal, L. T., Air Force Aeronautical Laboratories, Nonmetallic Materials Division, Mechanics and Surface Interactions Branch, AFWAL/MLBM, Wright-Patterson AFB, Ohio 45433.
- [10] Rich, M. J., University of Dayton Research Institute, Dayton, OH 45469.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PROCESS CONFIGURATION ALTERNATIVES FOR SEPARATION OF  
GAS MIXTURES BY PRESSURE SWING ADSORPTION

Prepared by:	Dr. Kent S. Knaebel
Academic Rank:	Assistant Professor
Department:	Chemical Engineering
University:	The Ohio State University
Research Location:	School of Aerospace Medicine Brooks Air Force Base Crew Technology Division
USAF Research:	Dr. Kenneth G. Ikels
Date:	August 29, 1984
Contract No.:	F49620-82-C-0035

PROCESS CONFIGURATION ALTERNATIVES FOR SEPARATION OF  
GAS MIXTURES BY PRESSURE SWING ADSORPTION

---

by

Kent S. Knaebel

ABSTRACT

Air separation has two important potential applications for military and commercial aircraft: to provide oxygen for life support of the crew and/or passengers, and to supply nitrogen as a means of repressurizing fuel tanks as they are depleted (or during descent). At present, technology is being developed by vendors to meet these needs with independent systems. The goal of the research described in this report is to establish the feasibility of conducting both of these separations in a common system. Results obtained by computer simulation indicate that the combined separation is feasible, and may be significantly more efficient than the independent systems, because essentially no by-products are released. Therefore, the amount of engine bleed-air that is necessary to extract both purified oxygen and nitrogen is minimized. Recommendations for future research are described.

## I. INTRODUCTION

There has been a strong effort in recent years to develop an on-board oxygen generation system (OBOGS), primarily for crew support in tactical aircraft. A parallel effort has been oriented toward an on-board inert gas generation system (OBIGGS), primarily for nitrogen blanketing of depleted fuel tanks. As the performance, i.e. product quality, reliability, etc., has improved it has become evident that a serious constraint of these systems is their demand for engine bleed-air.

Typically, either the OBOGS or OBIGGS will extract only a few percent of the incoming air as pure product, i.e. oxygen or nitrogen. The remainder, which is the bulk of the feed, is exhausted. As a result, other on-board systems that use engine bleed-air will be deprived, or engine modifications must be made that reduce its efficiency.

The ironic element of this situation is that the exhaust of an OBOGS is enriched in nitrogen, while the exhaust of an OBIGGS is enriched in oxygen. Thus, these waste products could be put to good use by merely interchanging and recycling them. That concept is the underlying principle of Complementary Pressure Swing Adsorption [1]. Recognition of this problem and its potential solution provided inducement for the research program, as explained in the next section.

## II. OBJECTIVES OF THE RESEARCH EFFORT

---

The principal goal of this research was to determine the practical feasibility of splitting air into its primary components, viz. oxygen and nitrogen, in a more efficient manner than by operating separate systems for each component. Complementary Pressure Swing Adsorption (CPSA) was suggested because this application was practically identical to one of the original cases for which the process was conceived. Other goals include: identification of future research needs for oxygen and nitrogen separation, which is described further in Section VII, and collaboration with Mr. John Kayser during his M.S.Ch.E. research, which is described in a separate report.

The concept of CPSA is simply to use the exhaust of one adsorbent bed to feed, purge, and/or pressurize the other adsorbent bed. The beds must exhibit "complementary" selectivity for the major components of the mixture. That is, the more strongly adsorbed component in the first bed is less strongly adsorbed in the second bed, and vice-versa. Since there is flexibility in the manner in which the exhaust products are recycled, there are many process layouts that might be attractive, depending upon the flow rates, compositions, adsorbate-adsorbent properties, pressures, temperatures, and economics. Three basic configurations of the CPSA process are shown in Figure 1 a, b, c.

Thus, a refined statement of the primary goal of this research is: to simulate several process configurations that are potential candidates for splitting oxygen and nitrogen from air, over a range of conditions that is suitable for aircraft applications.



### III. BACKGROUND

-----

Pressure Swing Adsorption (PSA) was developed by Skarstrom [2], and has been used primarily to purify one component from a mixture. There has been gradually increasing attention paid to research on this basic process recently because it is rugged, efficient, and effective for a variety of applications. Some of the recent research is reviewed elsewhere [3,4,5]. Mathematical modeling of the PSA process has advanced on two major fronts: analytical models developed by the method of characteristics, and numerical models that account for continuity, pressure losses, thermal effects, and mass transfer resistances by iterative solution of the governing differential equations. The former is relevant to the present research. Notable steps in that development have been made by Shendalman and Mitchell[6], Chan et al.[7], and Flores Fernandez and Kenney [5]. The present work is an extension of the work by Knaebel and Hill [1,8,9].

In the latter work, it was found that it is possible to extract simple analytical expressions for the molar quantities entering and leaving an adsorbent bed during the discrete steps of a pressure swing adsorption cycle. These expressions accounted for the effects of pressure, temperature bed porosity, adsorption equilibrium, and gas-phase composition. The inherent assumptions included: isotherm linearity, isothermality, and absence of all dissipative effects, such as diffusional resistance, pressure drop along the bed axis, and axial dispersion. Implementation of the concepts in a computer program is described in the next section.

#### IV. COMPUTER SIMULATION

---

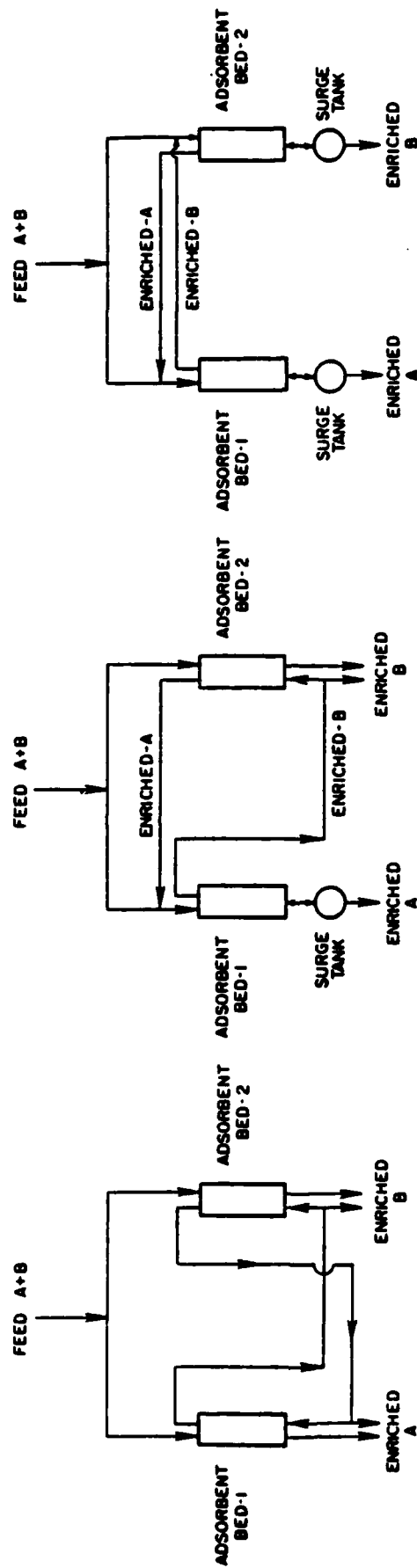
It was decided at the outset of the research effort to develop the computer simulator in as general a manner as possible. The inducement for this was to avoid time-consuming revisions during the study of various alternatives. As a result, the program was assembled in a modular structure with specific subprograms to accomplish repetitive tasks, and with the variables defined in arrays, rather than individually. Only the major concepts are covered in this section; the detailed flowchart and a sample of the Summary Output are in the Appendix.

The first step was to develop what might be considered as the "least common denominator" of a multiple bed adsorption system. The idea here was to enable a complex adsorption process to be decomposed into subsystems that could be easily and independently analyzed. The simplest subsystem was a single bed and the associated (real or hypothetical) tanks that accounted for capacity during stepwise filling or depletion. The primary considerations were to allow either the product or exhaust of any bed to be used for feeding, purging, and/or pressurizing another bed, and to allow the computer to simulate the operation sequentially. The "least common denominator" is illustrated in Figure 2.

Thus, in a particular cycle each bed (in a dual bed process) was simulated with a fixed feed composition (and other operating conditions) through its elementary steps (i.e. pressurization, high-pressure feed, and purge), and material was transferred

between the holding tanks at the end of the cycle. This prevented conflicts within the process during all cycles and permitted assessment of the time required for reaching steady-state from an arbitrary starting condition.

In comparing any one of the three configurations shown in Figure 1, with the elementary system shown in Figure 2, it is clear that several of the optional streams and tanks are unnecessary. On the other hand, by scanning all three configurations, while keeping the features of each in mind, it is easy to see that practically all of the features are necessary in order to simulate the entire set of three. The few additional features are included in anticipation of expansion to multicomponent mixtures and to accomodate more complex process configurations.



(a)

(b)

(c)

Figure 1. Basic complementary pressure swing adsorption flowsheets. Components A and B correspond to oxygen and nitrogen, respectively. Adsorbent-1 corresponds to 5A zeolite, and adsorbent-2 is 5A carbon sieve or 4A zeolite.

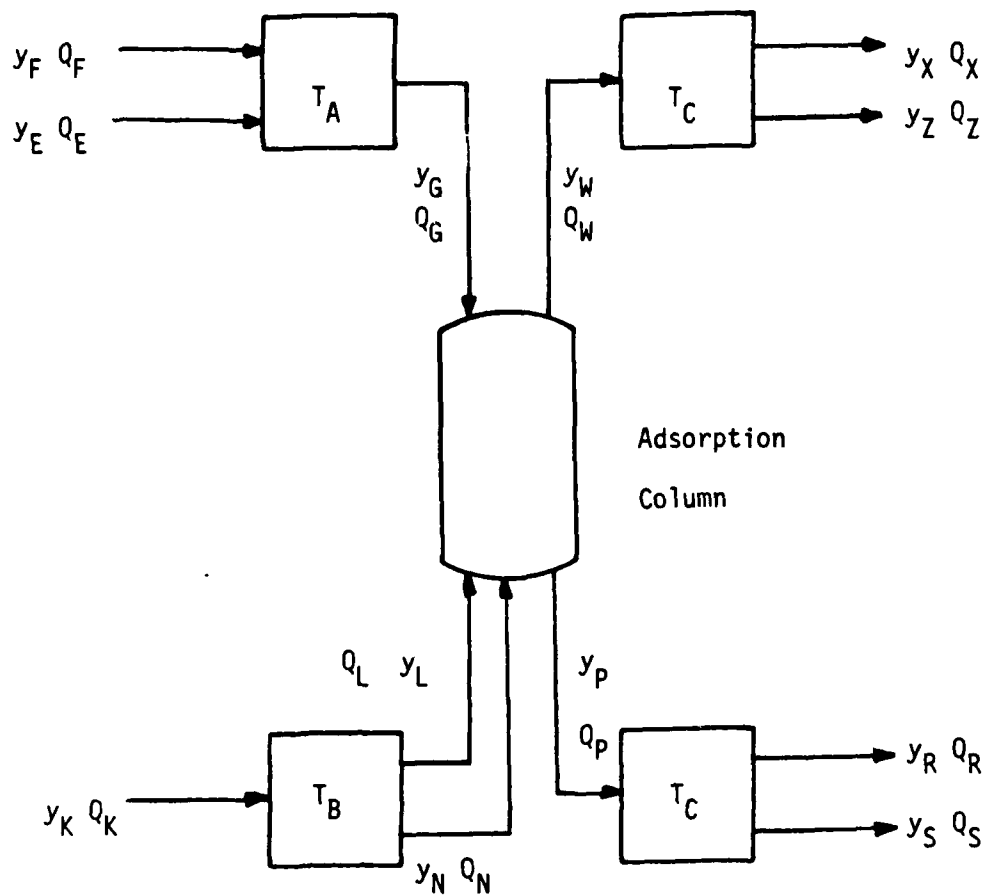


Figure 2. General flowsheet for a single adsorption column with associated tanks and streams. Note that the tanks and streams may or may not be present in a particular system.

## V. RESULTS

-----

Six case studies have been evaluated by computer simulation. They are listed below with numbers corresponding to each column within the system.

- A. Ordinary (Dual Column) Pressure Swing Adsorption:  
Product 1 to Purge 2 / Product 2 to Purge 1
- B. Single Column Pressure Swing Adsorption:  
Product 1 to Purge 1
- C. Complementary Pressure Swing Adsorption (see Figure 1a):  
Exhaust from 1 to Purge 2 / Exhaust from 2 to Purge 1
- D. Complementary Pressure Swing Adsorption (see Figure 1b):  
Exhaust from 1 to Purge 2 / Exhaust from 2 to Feed 1
- E. Complementary Pressure Swing Adsorption (see Figure 1c):  
Exhaust from 1 to Feed 2 / Exhaust from 2 to Feed 1  
(Numerical results are tabulated in the Appendix.)
- F. Complementary Pressure Swing Adsorption (reverse of D):  
Exhaust from 1 to Feed 2 / Exhaust from 2 to Purge 1

The only parameter that was systematically varied was the pressure ratio, which is the highest pressure (i.e. during the feed step) divided by the lowest (i.e. during the purge step). Other parameters that were fixed for all of the case studies were: adsorption isotherms, bed porosity, bed volume, and temperature. It was assumed that the adsorption of oxygen and nitrogen on zeolite 5A and on either zeolite 4A or carbonaceous sieve 5A, was complementary. That is, nitrogen was preferentially adsorbed by the zeolite 5A, and oxygen was preferentially adsorbed by zeolite 4A or carbonaceous sieve 5A. Furthermore, the relative preferences were taken to be symmetric. It is expected, however, that the actual preferences of the latter adsorbents for oxygen are of kinetic origin.

Case Study A. This system corresponds to one for which a direct, analytical solution exists and those results have been presented elsewhere [8]. It was possible to compare the steady state result obtained by the computer simulation with the results of simple algebraic equations. The results were identical, confirming the result of the computer model.

Case Study B. This system is directly analogous to Case A, and the direct, analytical solution was intuitively expected to apply to this configuration as well. It turned out that the results were identical, confirming intuition. A pragmatic result of this result is that it may be possible to simplify the traditional PSA system by using one column instead of two. This would eliminate some valves and their inherent synchronization problems.

Case Study C. This case involves the simultaneous solution of several coupled nonlinear equations, which inhibit a direct solution. For this case, unfortunately, the results were unattractive: there was only a very small range of pressure ratios (2 to 3) in which performance was acceptable. Thus, this configuration cannot be considered as a viable alternative for splitting oxygen and nitrogen.

Case Study D. This case resembles Case C, so a direct solution is not practical. The computer simulation was successful, however. Performance was acceptable at pressure ratios greater than 4, with the fractional byproduct (which is mainly comprised of nitrogen) rate diminishing and the byproduct purity increasing as the pressure ratio increased. At high pressure ratios

the ultimate byproduct flow is about 25% of the total feed to the process, and its molar purity is about 87% nitrogen.

Case Study E. This is a unique case because impure byproducts are unnecessary. That is, a binary mixture may, essentially, be totally separated into pure components. Performance was acceptable at pressure ratios greater than 2, and no fractional byproduct was produced. As a result, the system has only two products which are pure oxygen and pure nitrogen. The only effect of raising pressure beyond a pressure ratio of 2 is to gradually reduce the size of the equipment necessary to separate a given quantity of material. The relative flow rates are unaffected, and the product purities are unaffected. These results are stimulating because of the implications of nearly perfect efficiency and effectiveness for either or both major components of a mixture (see Appendix for results).

Case Study F. This case is essentially a mirror image of Case C, so a direct solution is not practical. In addition, the parameters were first adjusted so that the conventions were exactly reversed, and identical results were obtained, as expected. Subsequently, the conventions were restored, and only "mirror image" results from Case C were obtained. As before, the performance was acceptable at pressure ratios greater than 4, with the fractional byproduct (mainly comprised of nitrogen) rate diminishing and the byproduct purity increasing as the pressure ratio increased. At high pressure ratios the ultimate byproduct flow is about 25% of the total feed to the process, and its molar purity is about 87% nitrogen.



## VI. CONCLUSIONS

---

The most dramatic result of this research effort was that complete separations of the two major components of a mixture may be accomplished without necessarily producing any partially enriched byproducts. The process configuration that performs in this manner is shown in Figure 1 c.

The potential impact of that result is that air may be split into both essentially pure nitrogen and oxygen, with practically no waste. Therefore, both products may be produced on-board aircraft with minimal engine bleed-air. Presumably, if say only oxygen were needed, then the pure nitrogen could be discarded, yet the process would still waste no oxygen as is being done with current OBOGS units.

In addition, other case studies identified alternate process configurations that are also very efficient. Those configurations, however, produce a small quantity of byproduct, compared with the amount produced by ordinary pressure swing adsorption systems. Therefore, those configurations are likely to exhibit significant advantages in certain applications. For example, it may be possible to achieve a better combination of flow rates and pressures because the systems are not inherently symmetric.

## VII. RECOMMENDATIONS

---

As mentioned in Section VI, the most dramatic result of this research effort was that a pressure swing adsorption process can completely separate two components from a binary mixture without simultaneously producing any impure byproducts. In addition, there exist other similar process configurations that are also very efficient, which may offer significant advantages despite the fact that they produce small byproduct streams. These findings are, however, based on theory and require experimental verification.

Three phases of research for the separation of air into oxygen and nitrogen are envisaged: fundamental research to examine potential adsorbents that are selective for oxygen, applied experimental research of the process configurations and operating conditions that appear promising, and modeling of the process to account for kinetic effects, which strongly affect oxygen-selective sorption.

The first phase of research will involve a survey of commercial adsorbents, especially zeolite 4A and carbonaceous molecular sieves. The primary interest at that point will be the basic characteristics of the adsorbents. After that initial screening, the adsorbent isotherms, intraparticle diffusivities, and porosities will be evaluated experimentally. Based on that data, a selection will be made of a suitable adsorbent for the process.

The second phase of research will involve systematic experiments to establish the effects of operating conditions on the performance

of one or two process configurations. The operating conditions that are relevant are: pressure ratio, flow rates, temperature, and cycle times. In addition, it may be possible to test the effect of bed geometry on performance.

The third phase of research will involve development of a new mathematical model to simulate the effect of rapidly varying pressure on the performance of an adsorbent bed. In that situation there is expected a large effect of intraparticle diffusion resistance. In addition, there may be significant pressure drop over the length of the column, and axial dispersion may be important.

Future research may pursue improvements in the air separation system, explore other separations (such as isomers and organic homologs) for which Complementary Pressure Swing Adsorption may be effective, and examine the fundamental characteristics of broad classes of adsorbents for diverse separation applications.

#### REFERENCES

-----

1. Knaebel, K. S., "Analysis of Complementary Pressure Swing Adsorption," Fundamentals of Adsorption, A. L. Myers & G. Belfort, Eds., Engrg. Foundation, New York, 273-282, 1984.
2. Skarstrom, C. W., Ann. N. Y. Acad. Sci., 72, 751 (1959).
3. Wankat, P. C., Separ. Sci., 9, 85 (1974).
4. Wankat, P. C., Percolation Processes, A. E. Rodrigues & D. Tondeur, Eds., Sijthoff and Noordhoff, New York, 443, 1981.
5. Flores Fernandez, G. and C. N. Kenney, Chem. Eng. Sci., 38, 827-834 (1983).
6. Shendalman, L. H. and J. E. Mitchell, Chem. Eng. Sci., 27, 1449-1458 (1972).
7. Chan, Y. N., F. B. Hill, and Y. H. Wong, Chem. Eng. Sci., 36, 243-251 (1981).
8. Knaebel, K. S. and F. B. Hill, (submitted to Chem. Eng. Sci., 1984).
9. Knaebel, K. S. and F. B. Hill, Separ. Sci. and Tech., 18, 1193-1219 (1984).

#### ACKNOWLEDGEMENT

The author is grateful for the opportunity to participate in the research program sponsored by the Air Force Office of Scientific Research through the Southeastern Center for Electrical Engineering Education. The facilities, technical interaction, and hospitality offered by the School of Aerospace Medicine and the Crew Technology Division were outstanding. Furthermore, the guidance and enthusiasm for this research given by Dr. Kenneth Ikels are greatly appreciated.

In addition, it was especially worthwhile and convenient to be able concurrently to follow the research of Mr. John Kayser, who also participated in the summer research program, and is a M.S.Ch.E. candidate at the Ohio State University. His research also was made possible through the previously mentioned sponsors.

APPENDIX  
-----

COMPUTER SIMULATION DETAILS  
-----

Table A1. Computer Simulation Output for Case Study E.

Results for Column 1

CODE	NO	BETA	PHI	Pmax									
22	BEDS	2	0	30000	0	14025	100	00000					
P	QE	QF	QH	QR	QS	QT	QV	YV	YU	YV	YU		
1 00	0 0627	0 0000	0 0627	0 0627	0 0000	0 0000	0 0627	0 7900000	0 7900000	0 7900000	0 7900000		
1 26	0 1766	0 0000	0 1511	0 1511	-0 0431	0 0431	0 1766	0 5542872	0 0000000	0 0000000	0 4456128		
1 38	0 2223	0 0000	0 1649	0 1649	-0 0244	0 0244	0 2223	0 5260441	0 0000000	0 0000000	0 4739559		
2 00	0 2798	0 0000	0 1821	0 1821	-0 0003	0 0003	0 2798	0 5002517	0 0000000	0 0000000	0 4997483		
2 51	0 3074	0 0449	0 2039	0 2039	0 0119	0 0000	0 3404	0 5534751	0 0000000	0 0000000	0 5728918		
3 16	0 3415	0 1020	0 2312	0 2312	0 0271	0 0000	0 4164	0 5964468	0 0000000	0 0000000	0 6352801		
3 98	0 3845	0 1738	0 2657	0 2657	0 0462	0 0000	0 5121	0 6305804	0 0000000	0 0000000	0 6874772		
5 01	0 4386	0 2443	0 3090	0 3090	0 0703	0 0000	0 6327	0 6576937	0 0000000	0 0000000	0 7307270		
10 00	0 7003	0 7020	0 5189	0 5189	0 1866	0 0000	1 2159	0 7099268	0 0000000	0 0000000	0 3188798		
12 59	0 8365	0 9292	0 6279	0 6279	0 2470	0 0000	1 5186	0 7207208	0 0000000	0 0000000	0 8333024		
15 85	1 0076	1 2152	0 7650	0 7650	0 3230	0 0000	1 8998	0 7292948	0 0000000	0 0000000	0 6456429		
19 95	1 2230	1 5753	0 9377	0 9377	0 4188	0 0000	2 3796	0 7361053	0 0000000	0 0000000	0 3753249		
25 12	1 4943	2 0286	1 1551	1 1551	0 5293	0 0000	2 9837	0 7415152	0 0000000	0 0000000	0 8634498		
31 62	1 8337	2 5994	1 4287	1 4287	0 6910	0 0000	3 7441	0 7458123	0 0000001	0 0000000	0 8877732		
39 81	2 2656	3 3178	1 7732	1 7732	0 8820	0 0000	4 7015	0 7492257	0 0000001	0 0000001	0 8948193		
50 12	2 8068	4 2223	2 2069	2 2069	1 1224	0 0000	5 9068	0 7519370	0 0000002	0 0000001	0 8968423		
63 10	3 4881	5 3611	2 7529	2 7529	1 4231	0 0000	7 4241	0 7540907	0 0000003	0 0000002	0 9020472		
79 43	4 3458	6 7946	3 4403	3 4403	1 8042	0 0000	9 3343	0 7558014	0 0000000	0 0000000			

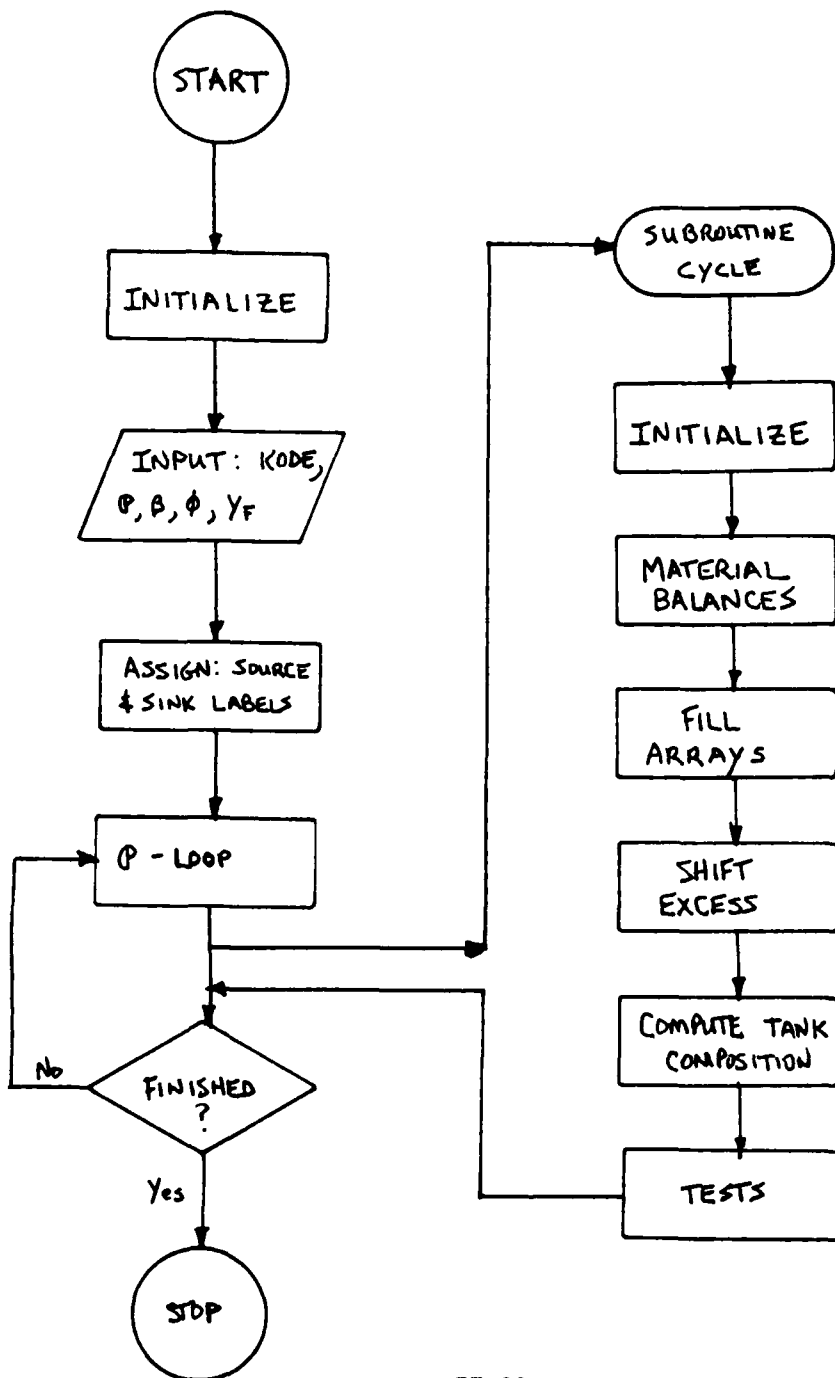
Results for Column 2

CODE	NO	BETA	PHI	Pmax									
22	BEDS	2	0	30000	0	14025	100	00000					
P	QE	QF	QH	QR	QS	QT	QV	YV	YU	YV	YU		
1 00	0 0627	0 0569	0 1196	0 1196	0 0000	0 0569	0 0627	0 7900000	0 7900000	0 7900000	0 7900000		
1 26	0 1766	0 0000	0 1511	0 1511	-0 0431	0 0431	0 1766	0 4456128	0 0000000	0 0000000	0 5342872		
1 38	0 2223	0 0000	0 1649	0 1649	-0 0244	0 0244	0 2223	0 4739559	0 0000000	0 0000000	0 5260441		
2 00	0 2798	0 0000	0 1821	0 1821	-0 0003	0 0003	0 2798	0 4997483	0 0000000	0 0000000	0 5002517		
2 51	0 3074	0 0119	0 2039	0 2039	0 0449	0 0000	0 3074	0 5002501	0 0000000	0 0000000	0 5189113		
3 16	0 3415	0 0271	0 2312	0 2312	0 1020	0 0000	0 3415	0 5447378	0 0000000	0 0000000	0 5386462		
3 98	0 3845	0 0462	0 2657	0 2657	0 1738	0 0000	0 3845	0 5939621	0 0000000	0 0000000	0 5850711		
5 01	0 4386	0 0703	0 3090	0 3090	0 2443	0 0000	0 4386	0 6366511	0 0000000	0 0000000	0 5799750		
10 00	0 7003	0 1866	0 5189	0 5189	0 7020	0 0000	0 7003	0 8130373	0 0000000	0 0000000	0 4337603		
12 59	0 8365	0 2470	0 6279	0 6279	0 9292	0 0000	0 8365	0 8312359	0 0000000	0 0000000	0 6360801		
15 85	1 0076	0 3230	0 7650	0 7650	1 2152	0 0000	1 0076	0 8441028	0 0000000	0 0000000	0 6756916		
19 95	1 2230	0 4188	0 9377	0 9377	1 5753	0 0000	1 2230	0 8543234	0 0000000	0 0000000	0 6832439		
25 12	1 4943	0 5293	1 1551	1 1551	2 0286	0 0000	1 4943	0 8624420	0 0000000	0 0000000	0 6953148		
31 62	1 8337	0 6910	1 4287	1 4287	2 5994	0 0000	1 8337	0 8689907	0 0000000	0 0000000	0 6946780		
39 81	2 2656	0 8820	1 7732	1 7732	3 3178	0 0000	2 2656	0 8740132	0 0000000	0 0000000	0 6989900		
50 12	2 8068	1 1224	2 2069	2 2069	4 2223	0 0000	2 8068	0 8780821	0 0000000	0 0000000			
63 10	3 4881	1 4231	2 7529	2 7529	5 3611	0 0000	3 4881	0 8813141	0 0000000	0 0000000			

Note: CODE refers to the simulated process configuration  
 BETA refers to the separation factor(isotherms etc.)  
 PHI refers to the column geometry, conditions, etc.  
 Pmax is the maximum pressure used in the simulation.

All other tabulated values are defined in Figure 2.

Figure A-1. Flowchart of Computer Program for Pressure Swing Adsorption Process Configuration Analysis.





1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE DISTRIBUTIONAL ANALYSIS OF CONTRAST SENSITIVITY MEASURES

Prepared by:	Dr. David L. Kohfeld
Academic Rank:	Professor
Department and University:	Department of Psychology Southern Illinois University at Edwardsville
Research Location:	Air Force Human Resources Laboratory Operations and Training Division Williams AFB, Arizona
USAF Research	Dr. Robert T. Nullmeyer
Date:	October 20, 1984
Contract No:	F49620-82-C-0035

# THE DISTRIBUTIONAL ANALYSIS OF CONTRAST SENSITIVITY MEASURES

by

David L. Kohfeld

## ABSTRACT

Visual contrast sensitivity measures were obtained using a video display that generated vertical sine-wave gratings. Threshold contrasts were determined using three different methods of stimulus presentation: Bekesy Tracking (Experiment 1), and the methods of adjustment and increasing contrast (Experiment 2). An inverted-U described the general form of contrast sensitivity functions for all three psychophysical methods, thus confirming that visual resolving power is best at intermediate spatial frequencies and progressively less keen at relatively extreme frequencies. The data from the Bekesy Tracking procedure were collected into separate distributions for the ascending and descending trials, and deconvolving the former distribution from the latter resulted in a model that was exponential in form. This analysis of component processes suggested that contrast sensitivity measures are comprised of at least two component processes, only one of which reflects sensory sensitivity, the other(s) having to do with undesired response-criterion bias. In general, the results indicate that the method of increasing contrast generates the most "criterion-free" measures of contrast sensitivity. These measures could serve as the best predictor of the contrast sensitivity required in other tasks, such as performance in a visual simulator, in which precise visual resolution is essential.

### Acknowledgements

The data collection, analyses and preparation of this paper were generously supported by the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education during the author's work as a Summer Research Fellow at the Human Resources Laboratory, Williams AFB, Arizona. Computer time and routines, including graphics, were funded by a research grant to the author from the Office of Research and Projects in the Graduate School at Southern Illinois University, Edwardsville. I am indebted to the research staff at AFHRL, Williams AFB, for their encouragement and support, including travel funds, without which this research could not have been accomplished. Special thanks are extended to my colleague, Dr. Robert Nullmeyer, for his close collaboration prior to and during this research assignment. Dr. George Geri provided the raw data for Experiment 1 and also offered incisive comments about the procedures used to generate and analyze contrast sensitivity functions. Finally, Mr. Glen York gave generously of his office space, computer-software expertise, and actual time to use his minicomputer terminal at AFHRL.

## I. INTRODUCTION

An important aspect of visual perception is the resolving power of the visual system, that is, its ability to determine very small spatial patterns, or to discriminate fine detail in a visual display. Visual acuity is particularly essential when a person is required to detect and identify objects and events in a visual environment where detailed information is sparse or fleeting. Some examples are driving an automobile at night, inspecting electronic microchips, or monitoring distant events as they are being displayed on a video screen. For the Air Force, the obvious example is pilot performance in either a visual simulator or the aircraft itself; objects and patterns that "approach" from a distance must be identified rapidly and accurately in order to accomplish a mission without adverse consequences.

One way to measure spatial resolution is to present a grating of vertical black test bars on a white background and to reduce the difference between the luminance of the light bars and the luminance of the dark bars until a person can no longer detect the dark bars on the white background. The coarseness of the grating can easily be varied, and the entire test field is presented on an ordinary video display. A crude example of this task is to consider the vertical black stripes on a zebra; if one could vary both the number and darkness of these stripes, what effect would this have on one's ability to detect the presence of stripes on this animal? Of course, if the recognition of "zebra" instead of "horse" were somehow important, this kind of visual resolving power would be essential. More specifically, the difference in luminance at which such striations are just resolved is the contrast threshold.

Contrast is usually defined as  $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$ , where  $L_{\max}$

is the luminance of the brightest part of the grating ("white stripes") and  $L_{\min}$  is the luminance of the darkest part ("black stripes"). The reciprocal of this contrast value is usually computed and is plotted as contrast sensitivity.<sup>1</sup> Contrast threshold measurements are usually obtained for a number of sinusoidal gratings of increasing spatial frequency, in cycles per degree (increasing number of stripes on the zebra). The resulting data are presented in the form of contrast sensitivity functions (CSFs), in which either contrast threshold<sup>2</sup> or contrast sensitivity<sup>3,4</sup> is plotted on a logarithmic ordinate against spatial frequency on a logarithmic abscissa.

Recent research by Ginsburg and his associates<sup>3,5,6</sup> has explored several psychophysical methods for obtaining contrast sensitivity measures, as well as demonstrating the relationship between CSFs and performance measures ranging from threshold detection and identification of letters to a pilot's skill in detecting a target during fighter-aircraft simulation. As evidence for the applied value of these procedures, Ginsburg and Cannon<sup>3</sup> also reported that visual abnormalities caused by brain lesions, multiple sclerosis, and cataracts are apparent in measures of threshold contrast sensitivity, but not necessarily in Snellen-type tests of visual acuity. Ginsburg consistently concluded that CSFs are superior to standard clinical tests (e.g., the Snellen test of visual acuity) in predicting performance in realistic visual environments.

## II. RESEARCH OBJECTIVES

The present research involved a distributional analysis of the data obtained with the three psychophysical methods studied by Ginsburg and Cannon; i.e., Bekey Tracking, the method of adjustment, and the method of increasing contrast. According to Ginsburg and Cannon, the method of

increasing contrast was the best of the three methods because the instructions to subjects were easiest to follow, less time was needed for data collection, the measures were most consistent (lowest variability), and perhaps most importantly in the present context, the contrast threshold values were least influenced by the response-criterion bias that obscures the "pure" relation between CSFs and sensory processes only. It should be noted, however, that none of the psychophysical methods provides data that are criterion free, primarily because an operator's disposition to respond to a detectable change in stimulation (i.e., contrast changes) tends to vary from trial to trial. In this vein, the separation of perceptual sensitivity and response bias from overall measures of threshold sensitivity is an enduring problem in signal detection theory and psychophysical measurement.

In the present investigation, entire distributions of contrast sensitivity measures were evaluated. A system identification method of deconvolution was used to decompose the distributions into component densities that presumably reflected the separate influence of "pure" contrast sensitivity and response bias on threshold detections. The general research objective was the same as Ginsburg's: to develop a valid predictor of contrast sensitivity in other tasks, such as performance in a visual simulator, in which precise visual resolution is essential.

### III. EXPERIMENT 1

A. Method. The operators (highly motivated, professional people) were three pilot trainees at Williams AFB, Arizona. Each operator had normal or fully-corrected vision. Testing was conducted in a darkened, quiet room in the Human Resources Laboratory.

The apparatus was an Optronix Model 200 Vision Tester, which has been

described elsewhere.<sup>3</sup> Briefly, this equipment is capable of displaying sinusoidal gratings at different spatial frequencies. Contrast at the video screen is defined as  $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$ . In the present experiment, the display screen had an average luminance of 100 cd/m<sup>2</sup>, or the equivalent of 87 decibels (51 millilamberts) of white light.<sup>7</sup> Seven spatial frequencies were employed: 0.5, 1.0, 2.0, 4.0, 8.0, 11.4, and 22.8 cycles per degree. The operators viewed the screen from a distance of three meters.

The psychophysical method employed in Experiment 1 is known as Beksy Tracking.<sup>2,3</sup> This method is available in a preprogrammed mode of operation in the Optronix Vision Tester. Specifically, the contrast either increased or decreased at a fixed rate of 4 dB per second, as determined by the position of the operator's hand-held switch. A block of trials was comprised of ten ascending and ten descending responses at a given spatial frequency. During a trial block, if a pattern on the screen was not detectable, the switch was held in one position until the grating was just barely visible (ascending), at which point the operator abruptly changed the switch direction until the grating seemed to disappear (descending). The Optronix microcomputer recorded a contrast threshold measurement whenever a change in switch position (threshold decision) occurred.

An experimental session was divided into five series of trial blocks, with a rest period between each series. Within a series, the seven programmed spatial frequencies were administered in order (0.5 to 22.8 cycles per degree). A total of 50 ascending measures and 50 descending measures for each of the seven spatial frequencies was obtained from each operator.

B. Results and Discussion. Figure 1 presents log threshold contrast sensitivity as a function of log spatial frequency, in cycles per degree,

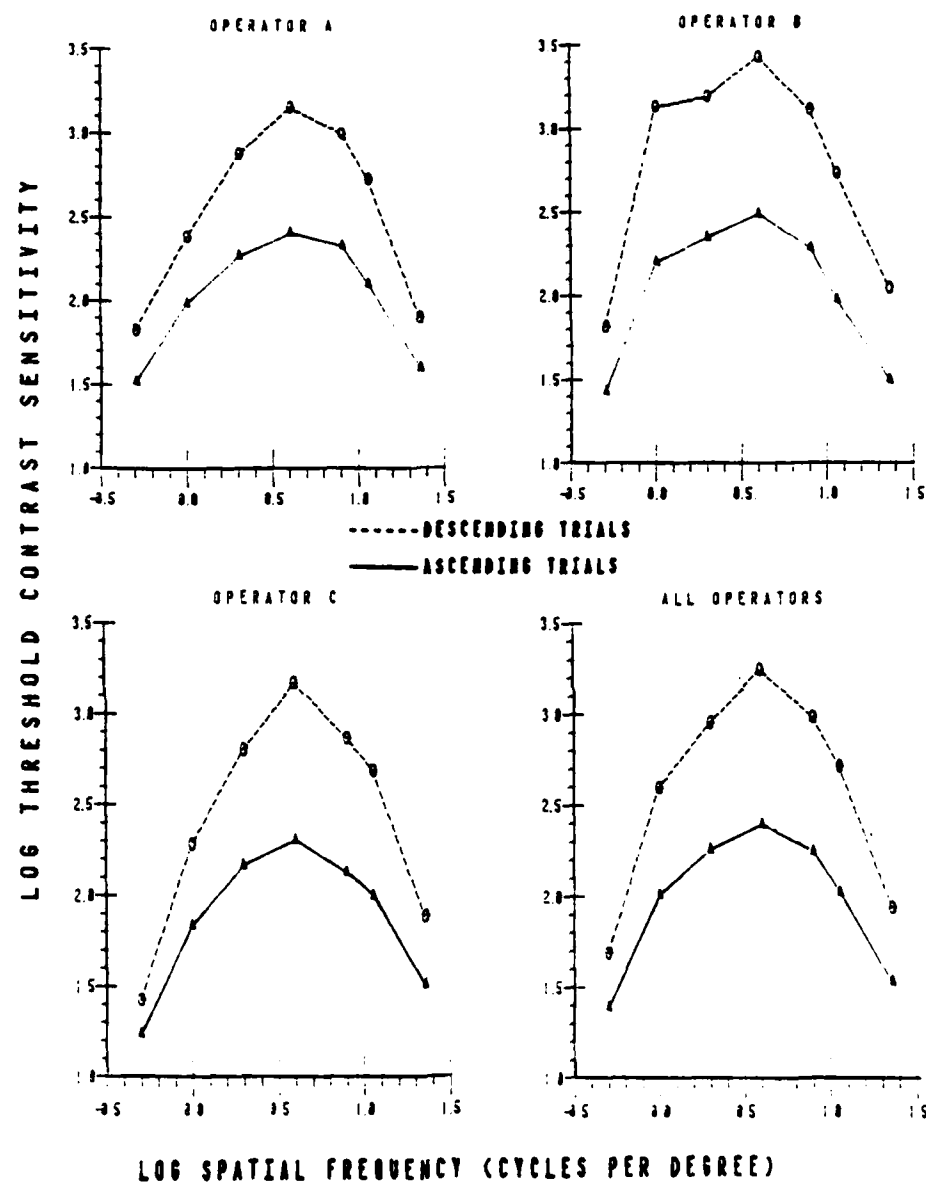


Figure 1. Contrast sensitivity functions for the Beksy Tracking method



for the Beksy Tracking method. Each operator's data are plotted separately (50 measures per data point), and the data means are pooled in the lower-right panel (150 measures per data point). The top function in each panel (dashed line) is for the descending trials (decreasing contrast), and the bottom function (solid line) is for the ascending trials (increasing contrast). The data form contrast sensitivity functions (CSFs) across the seven spatial frequencies (0.5, 1.0, 2.0, 4.0, 8.0, 11.4, and 22.8 cycles per degree) that are plotted on the logarithmic abscissa.

At least three conclusions may be drawn from the data shown in Figure 1. First, the CSFs reveal the characteristic inverted-U shape found in previous research.<sup>3,4</sup> This indicates that the resolving power of the visual system is best at intermediate frequencies (2-8 cycles per degree) and progressively less keen at relatively extreme frequencies (0.5 and 22.8 cycles per degree). Second, the CSFs for the descending trials suggest considerably greater contrast sensitivity than those for the ascending trials. This difference was also noted by Ginsburg and Cannon,<sup>3</sup> and is probably attributable to the exaggerated response bias that occurs when descending trials are coupled with ascending trials, as required by the Beksy Tracking procedure. This particular consideration will be evaluated subsequently more fully. Finally, the CSFs are both uniform and consistent across the three operators. However, this observation is based on graphic presentation of mean values, and deserves further evaluation based on the assessment of variability as well.

Figure 2 provides a detailed presentation of the same data that are shown in Figure 1. The data are now plotted as histograms (bin widths of 0.1 log sensitivity), with proportion of threshold detections on the ordinate as a function of log threshold contrast sensitivity on the

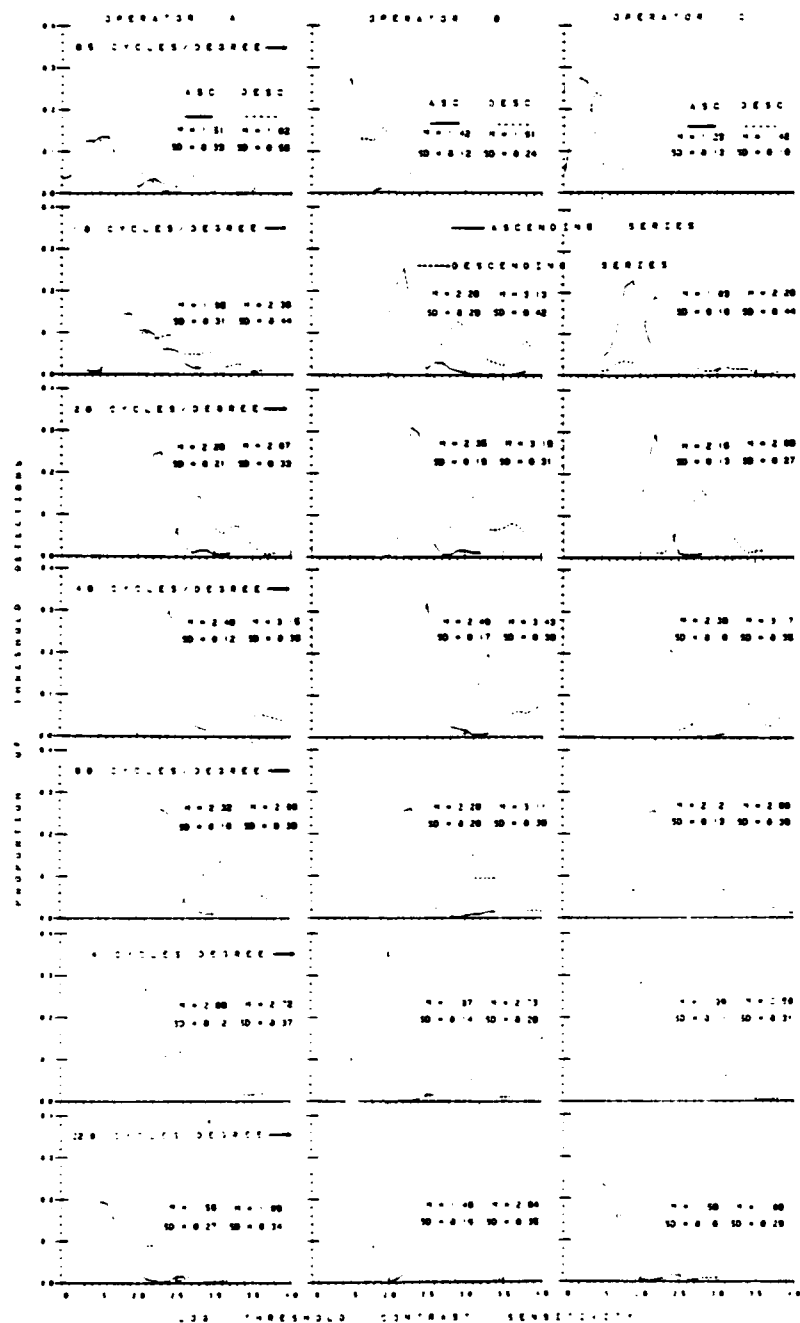


Figure 2. Individual distributions at each spatial frequency for the  
Bekesy Tracking method

abscissa. Figure 2 also shows the mean and standard deviation for the ascending and descending series for the three operators (rows) in each of the seven spatial frequencies (columns). It may be seen that in every instance the standard deviations are larger for the distributions of descending measures. In other words, although the operator's responses indicated higher mean contrast sensitivity in the descending (decreasing contrast) trials, the variability of these threshold measures was always greater.

The value of presenting entire distributions of data, as shown in Figure 2, is based on the assumption that the particular form of a histogram sometimes provides valuable information about the nature of the processes that give rise to a particular distribution of measures. Clearly, the shapes of the descending distributions are different than those of the ascending distributions, i.e., the former are generally flatter and more skewed than the latter. This raises the possibility that the operators were behaving differently in the ascending and descending modes. Perhaps the operators were perceiving the stimulus gratings differently, or, as suggested by Ginsburg and Cannon,<sup>3</sup> they may have adopted different response strategies. Based on these considerations, it seemed appropriate to pursue the matter of distributional analysis in more detail.

Although the individual histograms in Figure 2 are somewhat irregular, viewing down a column reveals that the shapes of the ascending and descending distributions are fairly consistent across spatial frequencies. Accordingly, the histograms were collapsed across spatial frequencies by adjusting the 50 raw scores in each distribution to a common origin of 1.0 on the log sensitivity scale, and then collecting these values across the seven spatial frequencies. Although this procedure eradicates infor-

mation about the actual positions of individual spatial frequency distributions on the contrast sensitivity axis, it does not sacrifice vital information about the nature of log sensitivity as a continuous random variable. The outcome is shown in Figure 3, which has the same axes and bin widths as in Figure 2. Thus, each operator's ascending and descending distribution in Figure 3 is comprised of 350 measures, and the pooled data of all three operators (lower-right panel) includes 1,050 measures per distribution.

In general, the ascending and descending histograms displayed in Figure 3 are most likely from the same Pearson "family" of distributions. That is, both distributions appear gamma in form; i.e., their means, standard deviations, and positive skewness are correlated. These properties suggest that both distributions were generated by common underlying processes that almost certainly have to do with the Békésy Tracking procedure for obtaining contrast sensitivity thresholds. But it is equally apparent that when descending thresholds are obtained, one or more of the component processes is behaving quite differently (larger mean, variance, and skewness), or perhaps another process emerges in addition to those required for ascending threshold estimations. In short, the shapes of the histograms in Figure 3 pointed to the feasibility of a deconvolution procedure which could reveal additional information about the processes that underlie contrast sensitivity distributions.

C. Distributional Analysis. The next step in this investigation was to identify and separate the component processes that characterize the difference between the ascending and the descending distributions shown in Figure 3. The convolution method employed for this task was the Z transform method<sup>8</sup> which is often associated with linear systems identifi-

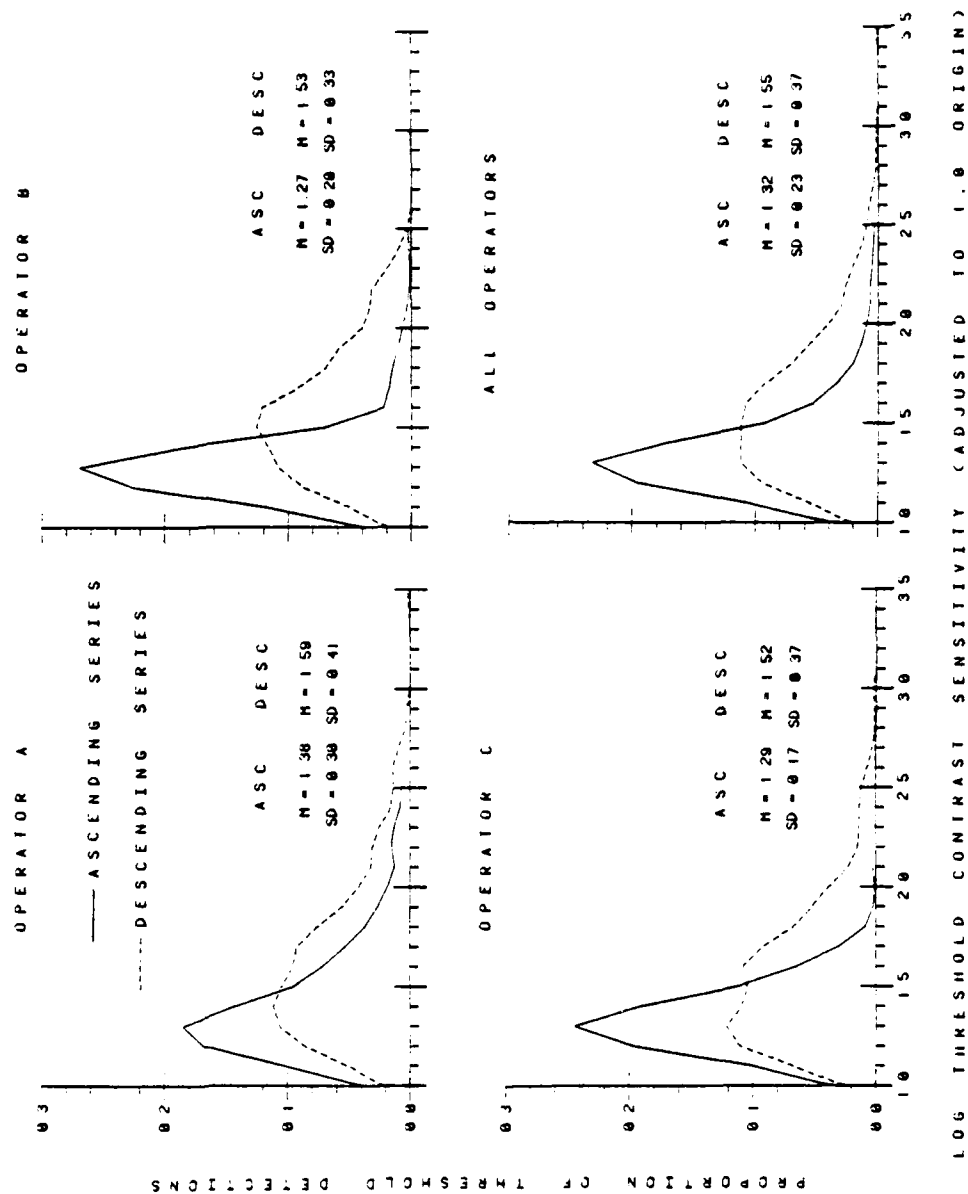


Figure 3. Individual distributions collapsed across spatial frequencies for the Beksy Tracking method

cation in engineering.<sup>9,10</sup> This approach incorporates the idea that an observed distribution (e.g., the descending series distribution) is often the convolution of a known distribution (e.g., the ascending series distribution) and another hypothetical process (e.g., a response bias component) for which one seeks to identify a model. Conceptually, convolution methods are similar to ordinary subtractive procedures except that the former involve the deconvolution (or "subtraction") of entire distributions rather than just measures of central tendency. Examples of the application of linear systems identification to the analysis of reaction time distributions are provided elsewhere.<sup>11,12,13,14</sup>

Using the Z transform method of convolution, a model of the difference between the ascending and the descending distributions in Figure 3 was obtained by deconvolving the ascending series distribution for all operators (lower-right panel) from their corresponding descending series distribution (same panel). The resulting difference model is shown in the middle panel of Figure 4. The ascending series distribution in the top panel of Figure 4 is the same as in the lower-right panel of Figure 3, and the distribution that is plotted as a solid line in the bottom panel of Figure 4 is the same as in the lower-right panel of Figure 3. The difference model selected by the convolution method was also the one whose reconvolution with the ascending distribution gave the closest approximation to the descending distribution from which it was identified. This reconvolution is the distribution composed of the dashed lines, as seen in the bottom panel of Figure 4. Although tests for goodness of fit were not made, the fit is obviously quite good, which supports the idea that the model describes a linear, additive component that accounts for the difference between the ascending and descending thresholds obtained in this research.

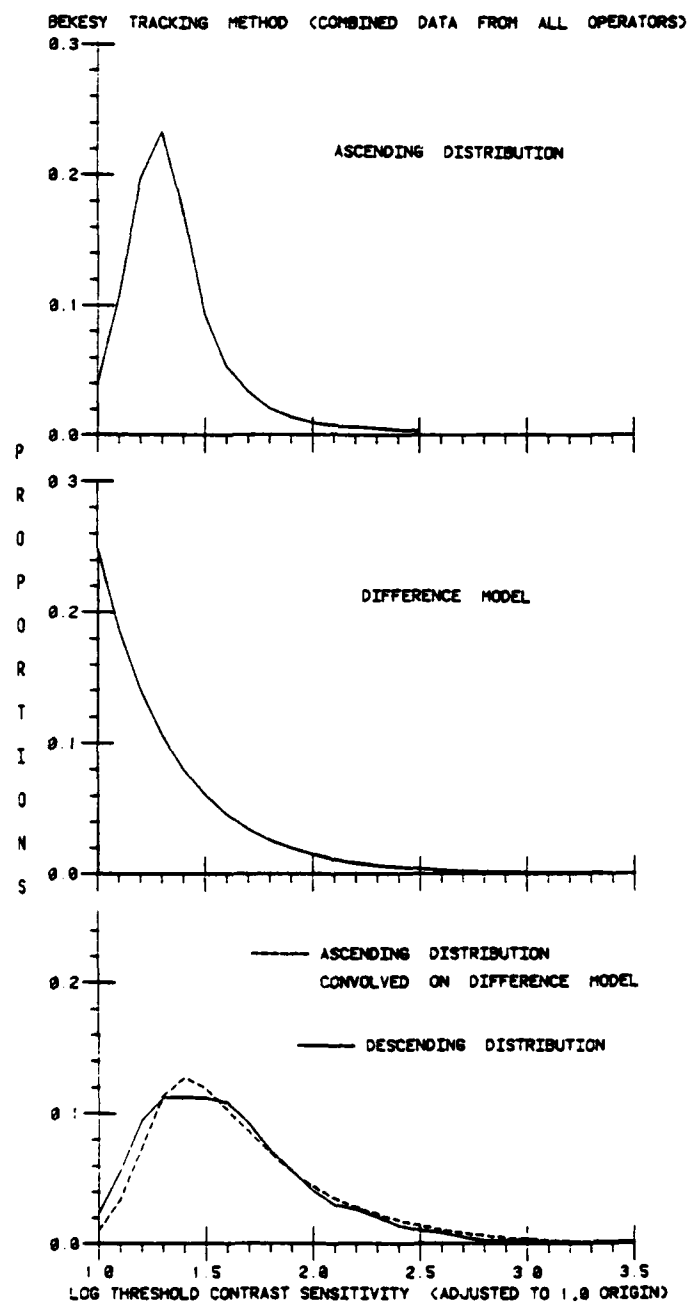


Figure 4. Difference model resulting from deconvolution of Bekesy ascending distribution from Bekesy descending distribution

The difference model that is shown in the middle panel of Figure 4 is exponential in form. An exponential model of the component process which separates the descending distribution from the ascending distribution implies that this underlying event operates much like a constant, conditional probability mechanism. Although this model is sharply bounded from below, the plotted location of this lower bound (1.0 log unit) on the contrast sensitivity axis is arbitrary, since the distributions from which it was derived were adjusted to a common 1.0 origin. The exponential form of this model, however, suggests a process that generates increasingly larger contrast sensitivity measures over a range of about 2.0 log sensitivity units. Moreover, the conditional probability that this process will influence the contrast threshold measures in the descending series becomes progressively lower according to a fixed probability rule. Finally, an exponential model provides a description of a single underlying event or process. As applied to the present analysis, this means that the higher and more variable contrast sensitivity thresholds in the descending series may be attributable to one categorical process.

#### IV. EXPERIMENT 2

The purpose of the second experiment was to obtain contrast threshold data for the method of increasing contrast and the method of adjustment. The goal was to compare these results with those obtained from the Békésy Tracking procedure, as was reported in Experiment 1.

A. Method. The operators were three Research Psychologists employed at the Human Resources Laboratory, Williams AFB, Arizona. All had normal or fully-corrected vision. DK, the principal investigator, was 43; a graduate student, MH, was 24; and a colleague, RN, was 34 years old. The three operators were familiar with the research aims and had extensive



experience with psychophysical methodology. The apparatus and general procedures were the same as in Experiment 1, except that only six spatial frequencies were utilized (0.5, 1.0, 3.0, 6.0, 11.4, and 22.8 cycles per degree).

The method of increasing contrast was utilized in the first experimental session. This method is essentially the same as the ascending series part of Békésy Tracking, except that no descending series are included. Briefly, the contrast increased at a fixed rate of 4 dB per second from below threshold to where the operator could first detect the presence of the test grating. The time between each threshold detection and the onset of the next trial was varied to discourage the operator from guessing the contrast threshold based on a fixed time interval from the last detection. The operator depressed a switch when the grating on the screen was just barely detectable. A block of trials consisted of 10 increasing presentations at each of the six spatial frequencies. There was a rest period at the end of each series of six trial blocks. As before, an experimental session was comprised of five series of trial blocks (six series for DK), yielding a total of 50 measures for each of the six spatial frequencies (60 for DK).

The method of adjustment was employed in a subsequent experimental session. The operator was given a hand-held switch with an adjustable knob that allowed him to either increase or decrease the contrast between the dark bars and the light bars on the screen until he could just barely detect the presence of the grating. The operator was instructed to turn the knob as far as it would go below the threshold before a trial began. He then increased the contrast and "bracketed" the threshold point, using as much time as was needed to achieve the desired setting. The operator

then depressed a button which stored the desired value in the microcomputer and reset the contrast on the screen to zero. All three operators reported making their threshold decisions quite rapidly, at a rate of about one every five seconds for the middle frequencies, but somewhat less rapidly at the extreme frequencies where threshold decisions seemed more difficult. As before, RN and MH generated 50 measures at each frequency, and DK recorded 60 values.

B. Results and Discussion. Figure 5 shows the mean threshold values, plotted in the same manner as in Figure 1. Because of the obvious similarity of the CSFs for each individual, the data were collapsed across operators, thus yielding 160 measures in each CSF data point for the adjustment method and increasing contrast method. The third CSF in Figure 5, labeled "Bekesy ascending trials," is the same as the one in the lower-right panel of Figure 1. This latter CSF is composed of data points with 150 measures, and is presented for comparison with the two CSFs obtained in the present experiment.

The most compelling feature of Figure 5 is the overall similarity of the CSFs for the three methods shown. This similarity is somewhat surprising when one considers that each CSF was composed of data taken from only three operators, and indeed, a separate trio of operators in the Bekesy ascending trials. These results suggest at least three important conclusions. First, the three different psychophysical methods generated such similar CSFs that one's confidence in the validity of these measures is improved considerably. Second, the elevated CSF for the Bekesy descending trials in Experiment 1 is probably based on properties of the method itself rather than "true" differences in contrast sensitivity. That is, the descending trials appear to be the "outliers" relative to three other

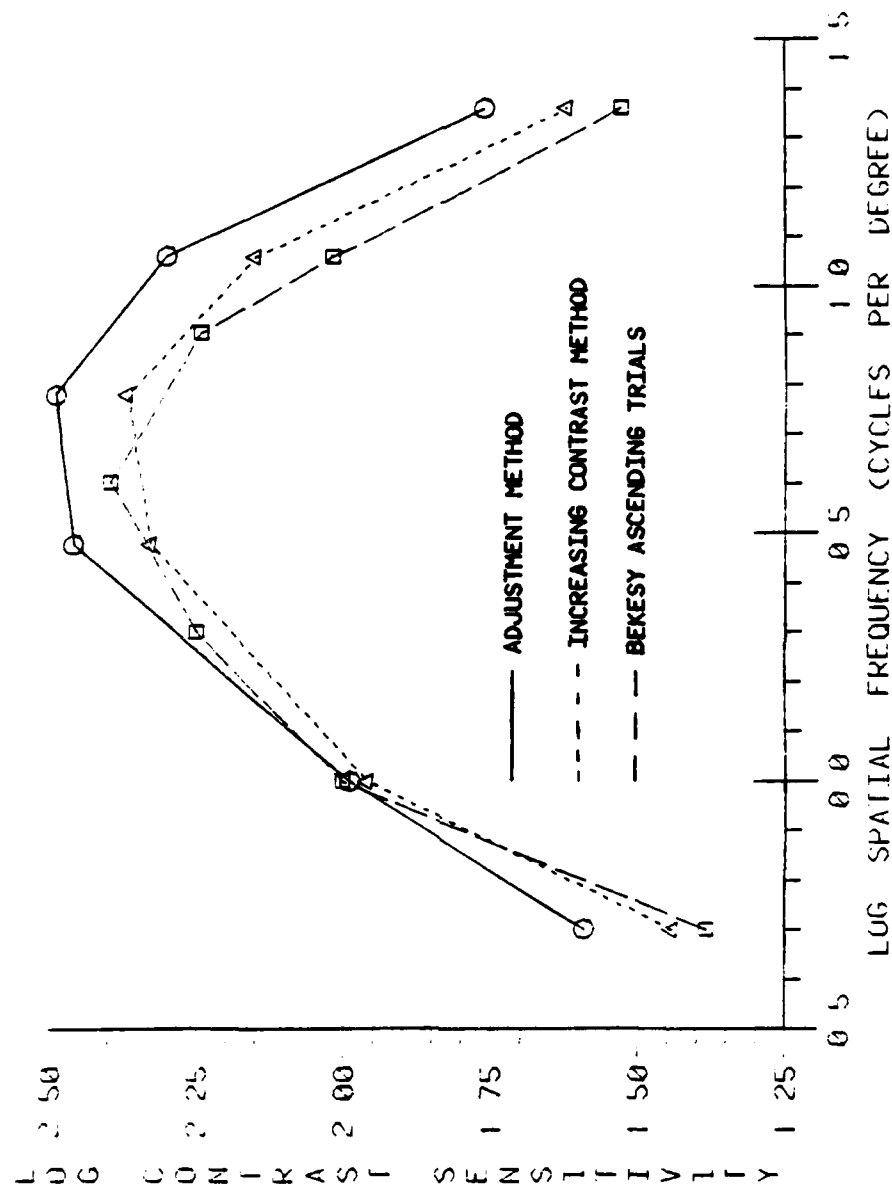


Figure 5. Contrast sensitivity functions for the adjustment and increasing contrast methods (Experiment 1) and for Bekey Tracking (Exper. 2)

procedures. Finally, it now seems possible that representative CSFs can be generated from relatively few operators, at least if these individuals are familiarized with the testing procedures and can generate a reasonably large sample of contrast threshold measures at each spatial frequency.

Figure 6 provides a comparison of the distributions of data for the same three methods that are depicted in Figure 5. The threshold values were adjusted to the 1.0 origin on the log sensitivity axis, just as they were in Experiment 1, Figure 3. The distribution composed of Bekesy ascending trials (1,050 measures) is the same as the one plotted in the lower-right panel in Figure 3, and in the top panel of Figure 4. Each of the distributions for the increasing contrast method and the adjustment method contain 960 values. The similarities between the general forms of the distributions shown in Figures 3 and 6 are quite apparent.

The results shown in Figure 6 lead to some important conclusions. First, all three distributions are gamma in form, with characteristic positive skewness. As noted previously, this implies that the processes or events which underlie each distribution are similar for the three psychophysical methods, as one might expect. Second, the Bekesy ascending trials and the adjustment method produced distributions that are similar in form, whereas the distribution of increasing contrast measures is more peaked, less variable, and less skewed. By implication, the perceptual and response processes which characterize the Bekesy ascending measures and the method of adjustment values should be similar, whereas the selective influence of at least one of these two processes on the threshold decisions obtained from the increasing contrast method should be reduced. Finally, if one accepts the view that a distribution of measures with the lowest possible variability is the most desirable in terms of measure-

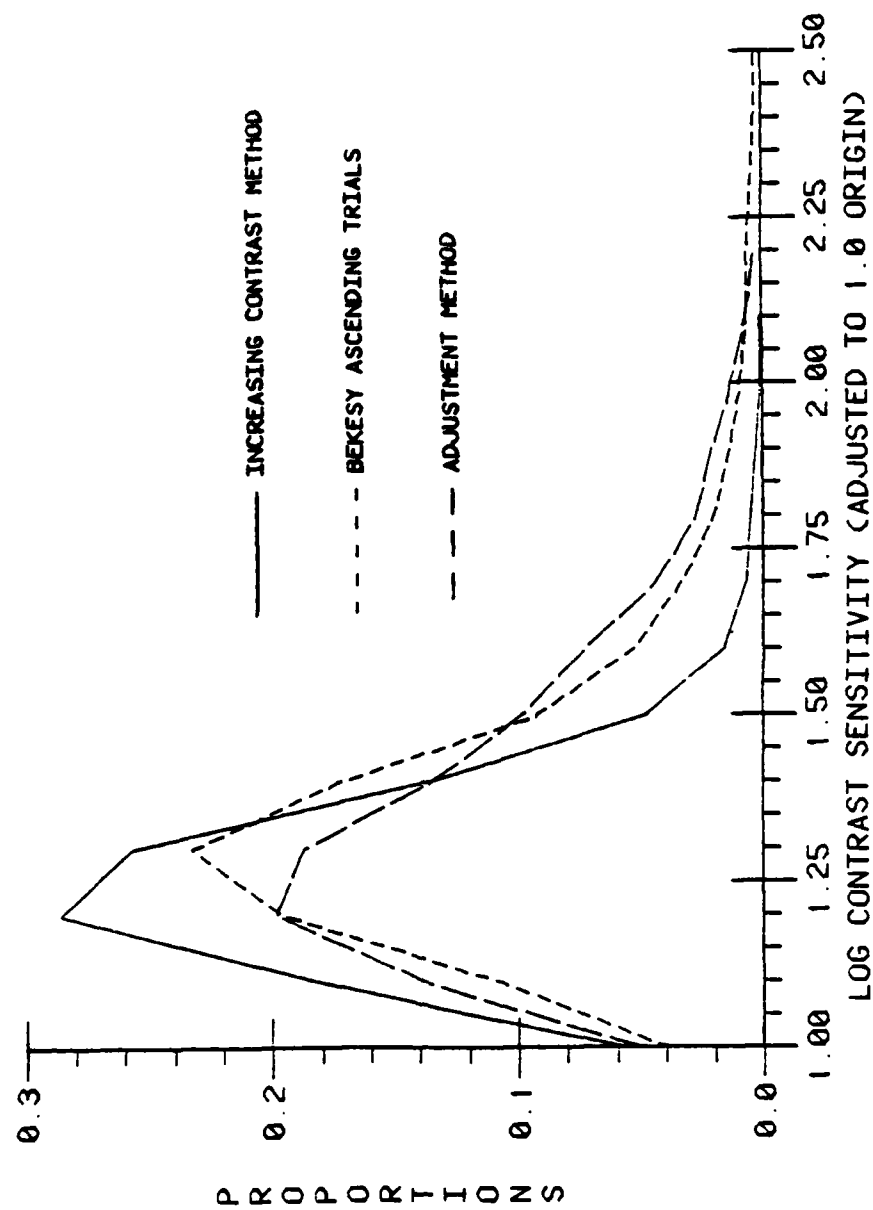


Figure 6. Distributions of the same data shown in Figure 5

mental precision, the method of increasing contrast could serve as the most predictive method for determining contrast sensitivity skills. Indeed, Ginsburg and Cannon<sup>3</sup> offered this conclusion for reasons quite similar to those presented in this report.

#### V. GENERAL DISCUSSION

In 1963, McGill<sup>15</sup> proposed that the form of some distributions would provide a kind of "signature" that reflects the nature of the processes that comprised these distributions. Since then, a number of investigators have demonstrated that simple reaction time is the convolution of two main components: sensory detection and response execution.<sup>11,14,16,17</sup> The present analysis of contrast sensitivity thresholds was influenced by this previous reaction time work in two important ways. First, the linear systems identification, and the associated Z transform method of convolution employed by Kohfeld, provided the analytic framework for decomposing the distributions of contrast sensitivity measures. Second, the reaction time process is sometimes discussed in terms of a response threshold that triggers the reaction; this is conceptually similar to the idea of the decision threshold that is determined with contrast sensitivity measures. In this vein, one can explore the parallels between the two reaction time stages, sensory detection and response execution, and the corresponding two stages in contrast sensitivity decisions.

The theory of signal detectability (TSD) provides one way to conceptualize the behavior of individuals in both reaction time and contrast sensitivity tasks. According to TSD, the operator is both a sensor and a decision maker, and therefore, the evaluation of threshold measurements would require an independent assessment of the operator's sensory sensitivity,  $d'$ , and response criterion,  $\beta$ .<sup>18,19</sup> TSD assumes further that  $d'$

and  $\beta$  are independent aspects of task performance, and are selectively influenced by different factors. With respect to contrast sensitivity, the visual contrast at the video screen determines  $d'$ , whereas  $\beta$  is characterized by the operator's strategy in responding to task demands. As such, the logic of TSD may prove useful in describing the forms of the distributions of threshold contrast sensitivity. Because of the nature of this threshold task, one can suppose that these distributions reflect the behavior of both sensory mechanisms ( $d'$ ) and a response process ( $\beta$ ). In other words, the distributions shown in Figures 2, 3, and 6 are a composite of at least two component densities,  $d'$  and  $\beta$ , each of which is an independent random variable.

It should be noted that no prior predictions were made about the shapes of the distributions actually obtained in this research. That is, no theoretical formulation dictated that particular forms of histograms should be generated, e.g., exponential, gamma, or normal. Thus, when the convolution method resulted in a model that was exponential in form (middle panel of Figure 4), no a priori logic suggested that this model was part of the sensory component ( $d'$ ), the response component ( $\beta$ ), or perhaps a third component altogether. However, both TSD logic and common intuition provide guidance in this matter. When one serves as an operator in the descending series part of the Beksy Tracking task, it is at once apparent that deciding when the grating has disappeared is considerably more difficult than when ascending threshold decisions are required. The consensus of the operators can be offered in the quote: "I held the switch down until I was sure that the bars were gone." This sounds very much like response bias, or in TSD terminology, the operators adopted a more lenient criterion in the descending trials, thus yielding inflated contrast thresholds.

These considerations lead one to conclude that deconvolving the ascending distribution from the descending distribution resulted in a difference model that reflects a response bias component. As noted previously, its exponential form implies that it represents a single process or categorical event. This interpretation does not imply that the ascending trials are free of response bias; rather, it supports the view that descending thresholds are influenced by at least two sources of undesired response bias. The separation of the  $d'$  and  $\beta$  components from the distributions generated with Békésy ascending trials, the method of adjustment, and the preferred method of increasing contrast (see Figure 6) is a matter for future research.

#### VI. RECOMMENDATIONS

The main goal of this investigation was to determine the psychophysical procedure and data analytic method that will provide the most valid measure of contrast sensitivity. The results suggest that the method of increasing contrast, coupled with the analysis of the forms of response distributions, is the recommended procedure for assessment of "pure" contrast sensitivity. An important implication is that the resulting measures will serve as a better predictor of the contrast sensitivity required in other tasks, such as performance in a visual simulator, in which precise visual resolution is essential. That is, when response bias is either reduced or analytically removed from contrast sensitivity measures, better prediction should result, primarily because the response bias in a specific predictor task (e.g., threshold determinations) will not necessarily correlate with the unique response bias in a different criterion task (e.g., identifying objects during simulated flight). Accordingly, two related lines of research are recommended at this juncture, one of which would deal with further



refinements of the contrast sensitivity measure itself, and the other having to do with using the measure as a predictor in an applied setting.

A. Refinements of Contrast Sensitivity Measures. There are a number of analytic procedures and different experimental approaches that deserve future research; three are mentioned here.

1. Hazard-function estimation. Considerable progress has already been made in estimating the hazard functions of the three distributions shown in Figure 6. While discussion of this work is beyond the scope of the report, it should be noted that all three distributions can be decomposed further by determining the form of their hazard functions. <sup>17,20,21,22</sup> These results suggest that all three distributions in Figure 6 represent the convolution of sensory and response components whose separate distributional forms can be estimated.

2. Experimental manipulation of task components. Established experimental procedures are available for selective experimental manipulation of the response bias component in measures of contrast sensitivity. The goal would be to provide empirical verification of the selective influences of sensory sensitivity ( $d'$ ) and response bias ( $\beta$ ) on these measures, which thus far in this investigation has been a matter of mathematical inference only.

3. Contrast discrimination thresholds. Almost all of the previous research with CSFs deals with measures of threshold detection. However, a more realistic task might include the operator's skill in discriminating different spatial frequencies. Consider an operator who continuously scans a visual display for new information. On the one hand, detection refers to the perceptual process of noting that "something" has appeared, disappeared, or changed, but this process is non-specific as to what it

was. On the other hand, discrimination is the more complicated process of identifying exactly what did appear, disappear, or change. Realistically, if a fighter pilot first detects something, subsequent identification of what it was (e.g., a "friendly aircraft" and not an "enemy fighter") may likely be the behavior one seeks to predict. In this vein, I have designed a series of experiments in which the operator's task would be to identify the particular visual frequency (number of vertical bars) as well as performing the preceding task of merely detecting that something had appeared on the video display.

B. Predicting Simulated Flight. An important research interest of Marc W. Hunter, who was assigned to this project as part of the GSSSP, was to explore the possibility that an operator's CSF could be used to predict the criterion skill of detecting and identifying objects during simulated flight. Previous research has shown that there is usually only moderate variation among different individuals' sensitivity to relatively high spatial frequency patterns (e.g., 22.8 cycles per degree). However, there are sometimes considerable individual differences in the low frequency range (e.g., 0.5 cycles per degree). It is to be noted that fog, haze, dust, and cloud cover has the effect of changing the spatial frequency of an object when introduced between some viewer and the object. Specifically, a haze or particulate will lower the spatial frequency of the object, having the effect of a "blurring" or spreading of the edges of the object by means of a property called atmospheric attenuation. Thus, for example, a pilot approaching some group of targets (e.g., aircraft) in a fog or light mist will usually first detect a group of masses that are nearly elliptical in shape, rather than perceiving them immediately as aircraft. If attenuation had not occurred the target aircraft would have appeared as being an object

having relatively high spatial frequency, with crisper lines showing more detail. Obviously, target aircraft will always be identified more quickly and accurately in a relatively clear atmosphere in which visual detail is keenest.

If contrast sensitivity measures had revealed significant differences in sensitivity to lower, and on occasion middle and upper spatial frequencies, it seems reasonable to assume that targets consisting of these frequencies would be detected and identified at differing distances. That is, those pilots with relative insensitivity to low spatial frequency patterns may require more time to detect targets composed of those low frequencies than those pilots shown to possess greater sensitivity in that specific frequency range. Obviously, the benefits of rapid and accurate target detection and identification are clear with respect to tactical air missions. Whether or not performance in a simulated environment would generalize to real-world environments is also a matter for further research.

Based on these considerations, Mr. Hunter explored the feasibility of a study to be performed at the Human Resources Laboratory on one of the aircraft simulation systems available there. It was determined that (a) the visual system currently used on the Advanced Simulator for Pilot Training (ASPT) has a setting regarded as "fog." However, this procedure merely changes the contrast of the scene produced. Since contrast change does not constitute spatial frequency change, it seems that the "fog" setting does not adequately simulate real fog effects precisely enough to manipulate it as the independent variable of primary interest: and (b) the Digital Image Generation System (DIGS) presently does not have the capability to produce atmospheric attenuation, and the necessary hardware and software changes were economically infeasible for a summer pilot study.

There is some indication that the Human Resources Laboratory is attempting to implement a visual system that could produce the necessary images for our recommended research. Even if this visual system did not allow for "flight" into the scene, the procurement of this system would allow research that would prove useful in determining the predictive relationship between contrast sensitivity functions and the detection of spatial frequency changes during simulated flight.

#### REFERENCES

1. J. P. Thomas, "Spatial Resolution and Spatial Interaction," in E. C. Carterette and M. P. Friedman (Eds.), Handbook of Perception V, New York, Academic Press, 1975.
2. R. Sekuler and P. Tynan, "Rapid Measurement of Contrast-Sensitivity Functions," American Journal of Optometry & Physiological Optics, Vol. 54, pp. 573-575, 1977.
3. A. P. Ginsburg and M. W. Cannon, "Comparison of Three Methods for Rapid Determination of Threshold Contrast Sensitivity," Investigative Ophthalmology & Visual Science, Vol. 24, pp. 798-802, 1983.
4. F. W. Campbell and J. G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," Journal of Physiology, Vol. 197, pp. 551-566, 1968.
5. A. P. Ginsburg, "Spatial Filtering and Vision: Implications for Normal and Abnormal Vision," in L. M. Proenza, J. M. Enoch, and A. Jampolsky (Eds.), Clinical Application of Visual Psychophysics, New York, Cambridge University Press, 1981.
6. A. P. Ginsburg, D. W. Evans, R. Sekuler, and S. A. Harp, "Contrast Sensitivity Predicts Pilots' Performance in Aircraft Simulators," American Journal of Optometry and Physiological Optics, Vol. 59, pp. 105-109, 1982.
7. D. L. Kohfeld, "Simple Reaction Time as a Function of Stimulus Intensity in Decibels of Light and Sound," Journal of Experimental Psychology, Vol. 88, pp. 251-257, 1971.
8. E. I. Jury, Theory and Application of the Z-Transform Method, Harrington, New York, Krieger, 1964.

9. T. C. Hsia, System Identification: Least Squares Methods, Lexington, Massachusetts, Heath, 1977.
10. A. P. Sage and J. L. Melse, System Identification, New York, Academic Press, 1971.
11. D. L. Kohfeld, J. L. Santee, and N. D. Wallace, "Loudness and Reaction Time: II. Identification of Detection Components at Different Intensities and Frequencies," Perception and Psychophysics, Vol. 29, pp. 550-562, 1981.
12. D. L. Kohfeld, "Stages of Reaction Time in Adults and Children," Paper Presented at the 22nd Annual Meeting of the Psychonomic Society, Philadelphia, Pennsylvania, November, 1981.
13. D. L. Kohfeld, "Reaction-Time Stages of Auditory Intensity," Paper Presented at the 23rd Annual Meeting of the Psychonomic Society, Minneapolis, Minnesota, November, 1982.
14. D. L. Kohfeld, "Sensory-Detection Models of Auditory and Visual Reaction Time Processes," Paper Presented at the 17th Annual Meeting of the Society for Mathematical Psychology, Chicago, Illinois, August, 1984.
15. W. J. McGill, "Stochastic Latency Mechanisms," in R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. 1, New York, Wiley, 1963.
16. D. M. Green and R. D. Luce, "Detection of Auditory Signals Presented at Random Times: III.," Perception and Psychophysics, Vol. 9, pp. 257-268, 1971.
17. S. L. Burbeck and R. D. Luce, "Evidence from Auditory Simple Reaction Times for Both Change and Level Detectors," Perception and Psychophysics, Vol. 32, pp. 117-123, 1982.

18. D. M. Green and J. A. Swets, Signal Detection Theory and Psychophysics, New York, Wiley, 1966.
19. W. N. Dember and J. S. Warm, Psychology of Perception (2nd Edition), New York; Holt, Rinehart, and Winston, 1979.
20. B. Bloxom, "Estimating an Unobserved Component of a Serial Response Time Model," Psychometrika, Vol. 44, pp. 473-484, 1979.
21. B. Bloxom, "Some Problems in Estimating Response Time Distributions," In H. Wainer and S. Messick (Eds.), Principles of Modern Psychological Measurement: A Festschrift in Honor of Frederick M. Lord, Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1983.
22. D. R. Miller and N. D. Singpurwalla, "Failure Rate Estimation Using Random Smoothing," Sankhya (Indian Journal of Statistics), Vol. 42B, pp. 217-228, 1980.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE RELATION BETWEEN HARD X-RAY BURSTS AND TYPE II RADIO EMISSION

Prepared by:	Gabriel Kojoian
Academic Rank:	Associate Professor
Department and University:	Department of Physics and Astronomy University of Wisconsin-Eau Claire
Research Location:	Air Force Geophysics Laboratory
USAF Research:	Dr. Ed Cliver
Date:	September 9, 1984
Contract No:	F49620-82-C09935



THE RELATION BETWEEN HARD X-RAY BURSTS  
AND TYPE II RADIO EMISSION

by

Gabriel Kojoian

ABSTRACT

The question of the relationship between the hard x-ray burst flares (HXRb) and the companion type II radio emission is investigated. An apparent positive correlation exists between the type II starting frequency (fundamental) and the delay between the peak intensity of the HXRb's and type II onset time. Suggestions for future research in this area are offered.

Acknowledgement

The author would like to thank the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him the opportunity to spend a scientifically interesting and worthwhile summer at the Air Force Geophysical Lab in Massachusetts. He would like to acknowledge the laboratory, in particular the PHP branch for its hospitality.

It is a pleasure to thank Dr. Ed Cliver for suggesting this area of research and for his collaboration and guidance. The author would also like to acknowledge the many helpful discussions with both Dr. Ed Cliver and Dr. S. Kahler.

## I. INTRODUCTION

It is well known that type II radio bursts are characterized by narrow, high frequency bands of intense radio emission that slowly drift towards lower frequencies. The drifting bands appear in harmonic pairs, i.e., type II events typically emit at a fundamental frequency and also have some emission at double the fundamental frequency. It appears that the basic emission mechanism is related to plasma oscillations. These oscillations are related to a characteristic frequency known as the plasma frequency. This frequency is one of the fundamental parameters of a plasma and depends on the plasma electron density. Since the coronal electron density appears to decrease with height, the lower the frequency of the radio emission the greater the altitude of its region of origin.

The type II radio bursts are manifested near or after the maximum of a chromospheric flare. These radio meter wavelength emissions first appear at the higher frequencies, i.e., about a few hundred megacycles per second, and then, drift downward to about 20 megacycles per second. The rate of this drift is on the order of one-fourth megacycle per second.

We have examined hard x-ray burst (HXRb) data collected by the Solar Maximum Mission satellite between 1980-82, and have obtained type II radio burst data for the corresponding period, from Solar Geophysical Data booklets. Our analysis of type II events associated with HXRb's indicate that the time delay between the peak of HXRb's and the start time of the meter-wavelength radio bursts was on the order of minutes. A graph of the type II starting frequency versus the time interval between the hard x-ray

and the corresponding type II radiation indicated an apparent positive correlation. The fundamental frequency of the type II bursts ranged from a lower limit of about 20 MHz to the neighborhood of 160 MHz, and the domain of the time delay extended from 0 to about 14 minutes. The analysis indicates an apparent cut-off of type II radio bursts with starting frequencies below 20 MHz. This frequency is in the neighborhood of the ionospheric cut-off frequency indicating a selective elimination of type II bursts with lower starting frequencies as observed from the earth. Thus this analysis suggests that type II radio bursts are more common than previously reported.

## II. RECOMMENDATIONS

- (1) We suggest a continuation of the study which suggests a positive correlation between the type II starting frequencies and the delay between the peak of the HXRB's and the onset time of the type II radio emissions.
- (2) We propose a study comparing the HXRB flares accompanied by type II radio bursts against HXRB flares without type II radio emissions. This may lead to a further understanding of the acceleration mechanisms operating on solar electrons. It is not known how the electrons are accelerated, nor is the source of the type II radio bursts known. A recently suggested concept is the "second-step" acceleration mechanism.<sup>(1)</sup> This may be tested with currently available data. If no significant difference exists between the HXRB spectra

with/without type II radiation, then type II events are not related to peak hard x-ray bursts, and the "second-step" acceleration mechanism is not supported. If there is a significant difference in the HXRB spectra with and without type II radiation, then at least two different acceleration mechanisms are acting and the "second-step" concept invoked to explain the acceleration of mildly relativistic electrons is supported. Thus, the results of this proposed study have the potential to either support or reject the "second-step" acceleration theory.

The HXRB data proposed for this study has been obtained by the Solar Maximum Mission satellite (during 1980-82). The type II radio burst information may be obtained from the Solar Geophysical Data booklets.

## REFERENCES

- (1) Bai, T., Ramaty, R., 1979, Ap.J., 227, 1072

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THERMAL LAYER DEVELOPMENT WITH ENERGY FLUX AT THE EARTH/AIR INTERFACE

Prepared by: Arthur A. Kovitz

Academic Rank: Professor

Department and  
University: Mechanical & Nuclear Engineering,  
Northwestern University  
Evanston, IL

Research Location: Air Force Weapons Laboratory, Nuclear Technology  
Office, Civil Engineering Research Division,  
Atmospheric Phenomenology Section (AFWL/NTEDA)  
Kirtland AFB, NM

USAF Research: Major Raymond L. Bell

Date: 1 September, 1984

Contract No.: F49620-82-C-0035

# THERMAL LAYER DEVELOPMENT WITH ENERGY FLUX AT THE GROUND/AIR INTERFACE

by

Arthur A. Kovitz

## ABSTRACT

The thermal layer at the ground/air interface, induced by blast radiation deposited in the ground, is studied using a one-dimensional version of the HULL code as the numerical tool. Energy flux at the interface is the sole agent causing disturbances in the air above. The calculations are qualitatively consistent with the results of others, showing weak thermal coupling between ground and air (a thermal layer no greater than 2m thick). Calculations also reveal the dependence of the results on cell size, and suggest several possible paths for determining the "correct" cell size. Work should continue on establishing a method for rationally eliminating cell-size dependence within the framework of the HULL, one-dimensional code written for this analysis. The ultimate goal is to provide a thermal layer segment for HULL that can be valid over a broad spectrum of events; i.e., that does not depend primarily on experimental correlations.



#### ACKNOWLEDGMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force Weapons Laboratory, Kirtland AFB, N.M. He would like to acknowledge, in particular, the Civil Engineering Research Division, Technology Branch, Atmospheric Phenomenology Section for its hospitality and excellent working conditions.

As my "EFFORT FOCAL POINT" Major Raymond L. Bell was a continuous source of ideas; he suggested the area of research, and introduced me to the world of the HULL code. It is also a pleasure to acknowledge the ever present, knowledgible support provided by Lt. Glenn James. The presence of Lt. Robert Connell gave the author many opportunities for helpful discussions. Mr. Lester Bowers, and Lt. William Kitch smoothed the author's way through the local maze of computers.

## I. INTRODUCTION

As an individual with a background in fluid mechanics and thermodynamics the opportunity to work on a problem embracing those areas was very attractive. It also was of interest to me that numerical methods would be stressed; my experience with hydrodynamic codes was limited. The concerns of AFWL/NTEDA focused on developing a model for the thermal layer that avoided experimental correlations, and could be appended to the HULL hydrodynamic code. These tasks compliment both my experience, and my desire to learn more about numerical methods in fluid mechanics.

## II. OBJECTIVES OF THE RESEARCH EFFORT

In March 1984 I visited Major Raymond L. Bell at AFWL/NTEDA, Kirtland AFB, NM. We discussed our *summer* research goals. It was decided that I should work on developing a physically based (as opposed to one based on experimental correlations) model for the thermal layer and blast wave precursor, and then incorporate it into the HULL hydrodynamic code. These goals remain; however, the emphasis this summer has concentrated on a one-dimensional version of the problem. This afforded me an opportunity to write my own one-dimensional HULL code (with much valuable help from Major Bell), and begin a process of evolving a thermal layer model.

## III. BACKGROUND

Nuclear blast precursors have been observed since the early days of testing. These precursors take the form of shock fronts moving at ground level ahead of the main blast wave. They induce in their wake a complex

flow structure that results in important modifications to the ideal blast flow. The implications of these precursor flows on the design of blast resistant objects gives them practical as-well-as scientific interest.

The existence of precursor flows is due to the formation of a heated layer of air at ground level prior to the arrival of the main blast effects. This thermal layer is produced by radiation energy from the blast source being deposited in the air-surface interface. The subsequent energy transfer to the adjacent air produces the thermal layer. This model assumes that the air is transparent to the incident radiation with adsorption occurring in the earth. Energy transfer to the air is then accomplished by several mechanisms, including conduction, and surface motion accompanied by turbulent mixing of hot and cold material. The resulting temperature increase in the air causes it to move in accordance with the equations of hydrodynamics.

This work is an attempt to describe the air motion, and thermal layer development as a function of a prescribed energy flux into the air at the air-surface interface. The treatment is that of a one-dimensional (time dependent) initial value problem using the HULL hydrodynamic code to integrate the governing Euler equations of motion, and the energy equation. Surface temperature as a function of time rises to an assumed upper limit equal to the temperature at which surface phase change occurs; it is held constant until the decreasing energy flux and energy loss mechanisms reduce surface temperature below this value.

This treatment ignores mechanisms of energy transport associated with mixing of different materials, and energy absorption in the air; i.e., dust entrainment, and energy absorption by dust. Its focus is on fluid motion in response to an energy flux. Extensions of this work to include

real fluid effects (viscosity, thermal conductivity, turbulent mixing, etc.) are appropriate future projects.

#### IV. BRIEF REVIEW OF THE LITERATURE

This section is intended to give an overview of reports and papers that played an immediate role in the research. As such it will not be a comprehensive review of thermal precursor literature.

Since the effort focused on two goals, (a) to develop a model of the thermal precursor, and (b) to incorporate that model into the HULL code, literature associated with those two areas will be noted.

A useful review [1] of the physics of dust-entrained shock waves describes the pre-shock thermal layer formation as a sequence of events involving several distinct processes. Among them are 1) surface and sub-surface heating, 2) conduction to the air, 3) blowoff of material from the surface, 4) convection due to the instability of the heated air layer, and finally 5) transfer of energy to the air by the material ejected from the surface. The last process is, probably, the primary agent for rapid energy transfer to the developing thermal layer. Further discussion of these events may be found in Ref. [2].

An introduction to the subject, that is intended for those not in the field, is given by Rosenblatt [3]. His report establishes a framework for the whole area of nuclear dust/debris phenomena.

The thermal layer model is to be incorporated into the hydrodynamics of the overall blast/shock-wave-induced flowfield. This requires the use of numerical techniques to obtain solutions to systems of equations describing the fluid physics. The HULL hydrodynamics computer code [4,5]

is one of several large algorithms for numerical solution of the hydrodynamics equations. It is particularly designed to compute the response of deformable media to large and rapid deposition of energy. A thermal layer model based on the results of Needham and Crepeau [6] has already been incorporated into the code. Although results are physically reasonable, the model is based on a fit to experimental data. The current effort seeks a thermal layer description less dependent on experimental input. More will be said of the Needham/Crepeau "Flux Dependent Thermal Layer Model" [6] when comparisons with the current model are made.

A comprehensive thermal layer model, with few ad hoc assumptions, has been constructed by Sandford, et al [7]. Its object is to describe the interaction between dust and radiation during a period of surface fluence. Computed results compare favorably with experiments conducted with a solar furnace and a "thermal radiation simulator". The observed phenomenology is closely related to the parameters of the calculation; of particular interest is the energetics associated with the dust/radiation interaction. Although this model is strongly related to basic processes, its complexity makes its use with general hydrodynamic codes difficult.

#### V. A ONE-DIMENSIONAL VERSION OF THE HULL CODE

The HULL code is a finite difference solution to the equations of continuum mechanics; see Ref. [8]. It exists in several versions. The one in use at this facility comprises about 80,000 lines of code. It is designed for universality of use, and generality. Application to a geometrically simple configuration, with the goal of examining some of the properties of the solution in the light of various assumptions, is best

done with a pared down version of the code. Accordingly, a strictly one-dimensional HULL code was written for this study. It is small enough to be run on machines such as the DEC 11/24 and/or the TEXTRONICS 4052, which affords convenient graphics capability.

The features of the HULL code are fully described in Refs. [4,5,8]. In briefest outline the solution path is as follows:

1. cells (EULERIAN) are defined along the x-axis in terms of cell-centers,  $XC(i)$ , and cell boundaries (left- and right-hand)  $X(i-1)$ ,  $X(i)$  such that
$$XC(i) = [X(i) + X(i+1)]/2$$
2. at time  $T(n)$  flow properties, and their spacial derivatives are known at  $X(i)$ ,  $XC(i)$ , and  $X(i+1)$ .
3. equations for the material time derivatives of all the flow properties are expressed in terms of finite difference analogs of the spacial derivatives appearing on their right-hand-sides.
4. values of flow properties at the LAGRANGIAN positions  $X(i)$ , and  $X(i+1)$  are obtained at  $T(n+1/2)$  for the half time-step  $dT/2$  through direct use of the material derivatives noted above.
5. these values are used to obtain updated, to  $T(n+1)$ , expressions for the cell-centered spacial derivatives; in effect, both cell- and time-centered values of the material time derivatives are now known.
6. the last step is a "REMAP" to obtain the corresponding EULEARIAN cell-centered properties; one now knows the flow properties at  $XC(i)$  at  $T(n+1)$ .
7. steps 1 through 6 are repeated, as required. Cell-size, and time-step are constrained by the need for computational stability, as expressed by the COURANT condition, i.e., the time-step must be

smaller than the time required for an acoustic signal to traverse any one of the cells.

8. boundary conditions are imposed by introducing "guard-cells" on the boundaries, exterior to the flow region, such that flow properties at the boundary satisfy the required values.

Working programs were written both in FORTRAN (for use with a DEC 11/24), and in BASIC (for use with a TEK 4052). It was verified that numerical results from the two codes were the same.

#### VI. A THERMAL LAYER MODEL

The one-dimensional HULL code was applied to an idealization of the thermal layer problem. An ideal gas in the half-space  $X > 0$  is energized by a flux ( $\text{ergs/cm}^2/\text{s}$ ) at  $X=0$ ; one expects the resultant local temperature rise, with zero flow velocity at  $X=0$ , to cause an expansion of the gas and motion to the right ( $X > 0$ ). This, in turn, exerts flow work on neighboring fluid particles, causing them to move and transmit energy to their neighbors. The process distributes energy into the region  $X > 0$  causing the formation of a "thermal layer" accompanied by pressure, density, and velocity disturbances.

HULL code hydrodynamics does not include energy transfer by thermal conduction. Therefore, the energy flux at  $X=0$ , with no mass flux there, cannot enter the gas, either through molecular collisions or convection; it must be assumed to be added "directly". Since the physics of the energy addition have been avoided there is a loss of information in the model. One may expect the calculated quantities to be, to a certain extent,

indeterminate. It will be seen that this is the case; the distributions of temperature, pressure, and velocity depend upon the chosen cell size.

Thermal conduction is a relatively slow energy transmission process; the observed rates of energy transfer [6] are much greater. This suggests that the inclusion of thermal conductivity would not be an important factor in reproducing the thermal layer as produced by blast. Other energy transfer mechanisms must be dominant. Among them is likely to be transfer by material ejected from the surface [1,2,3,6]. Ultimately, thermal conduction must play the final role; but the increase in surface area, and interpenetration of transfer surfaces results in the required enhancement of the energy transfer rate. A more realistic model should include material injection and subsequent energy transfer.

Energy transfer by material injection is probably confined to a gas layer close to the interface  $X=0$ . The thickness of this layer (it relates to the cell size in a HULL code calculation) is determined by the physics of energy absorption in the ground, and mixing of ejected material with the overlaying medium. Above this layer rapid energy transfer occurs by flow work only, consistent with the HULL calculation. This picture of the thermal layer development admits the "direct" energy transfer (by material injection, actually) into the gas layer contiguous with the surface; it does not, however, determine the thickness of that layer.

In this work some calculations are completed showing qualitatively correct thermal layer behavior with energy flux at  $X=0$  deposited into the HULL cell contiguous with the ground/air interface. There is, as noted above, a dependence on cell size. These results will be discussed more fully in the next section.

Needham and Crepeau's [6] model introduces experimentally observed



temperatures both at ground level and above, together with a threshold flux concept, to obtain results in good agreement with many other experiments. Their calculations avoid the indeterminacy of cell size by incorporating the actual physics of the energy transfer through the use of experimental observation. In so far as their generalizations of observation are correct their method will have widespread applicability. The ultimate goal of this effort is to be able to do the same with a reduced need for a priori experimental input.

## VII. MATHEMATICAL FORMULATION

The governing equations for one-dimensional flow of a compressible fluid with negligible viscous forces and negligible heat flux internal to the medium are

$$D\rho/Dt = - \rho \partial u / \partial x \quad (1)$$

$$Du/Dt = - (\partial p / \partial x) / \rho \quad (2)$$

$$DP/Dt = - \rho c^2 \partial u / \partial x \quad (3)$$

$$DE/Dt = - (\partial Pu / \partial x) / \rho \quad (4)$$

In these equations

$\rho$  is the density; gm/cm<sup>3</sup>

$u$  is the fluid velocity; cm/s

$P$  is the pressure; dynes/cm<sup>2</sup>

$c(\rho, P)$  is the local sound speed; cm/s

$E = I(\rho, P) + u^2/2$  is the total energy; ergs/gm

$I(\rho, P)$  is the internal energy; ergs/gm.

The quantities  $D\rho/Dt$ ,  $Du/Dt$ ,  $DP/Dt$ , and  $DE/Dt$  are material time derivatives;  $\partial U/\partial x$ ,  $\partial P/\partial X$ , and  $\partial(PU)/\partial X$  are partial derivatives with respect to the spacial dimension  $X$ . Knowledge of  $\rho$ ,  $u$ ,  $P$ ,  $c$ , and the spacial derivatives of  $u$  and  $P$  at any instant determine the material derivatives. Since the internal energy  $I$  is a function of  $P$  and  $\rho$ , Eqs. (3,4) are not independent. Convenience will indicate which one to use.

For a chosen medium the equation of state  $P(\rho, T)$ , the "caloric" equation of state  $I(\rho, P)$ , and  $c(\rho, P)$  are determined;  $T$  is the absolute temperature, in degrees K. This work assumes the ideal gas equation of state with constant specific heats;  $\gamma$  is the specific heat ratio. Then

$$P = \rho RT, I = P/[\rho(\gamma-1)] \quad , \quad c^2 = \gamma P/\rho \quad . \quad (5)$$

The boundary and initial conditions in terms of the independent variables  $(X, t)$  are  $u(X, 0) = 0$ ,  $u(0, t) = 0$ ;  $\rho$ ,  $P$ ,  $T$ , and  $I$  are their uniform, atmospheric values at  $t=0$ .

Energy input at  $X=0$  is a specified function of time. In this case a flux  $q(t)$  is chosen such that it approximates one used by Needham and Crepeau [6];

$$q(t) = Q_m(t/t_m)\exp[-B(t-t_m)^2] - Q_e(T_l - T_a)/T_l \quad ; \quad (6)$$

in Eq. (6)  $Q_m$  is the maximum flux,  $t_m$  is the time to maximum flux,  $B$  is a constant that depends upon  $t_e$ , a time such that  $Q$  has decreased to  $Q_e$ ;  $T_1 = T(0,t)$ , and  $T_a$  is the atmospheric temperature. The term  $Q_e(T_1 - T_a)/T_1$  represents a loss mechanism at  $X=0$ . It causes the flow field to relax back to its quiescent state for  $t \gg t_m$ ;  $q(0) = 0$ ,  $q(t \gg t_m) \rightarrow 0$  for  $T_1 \rightarrow T_a$ . In this application

$$Q_m = 7.83 \times 10^9 \text{ ergs/cm}^2/\text{s},$$

$$Q_e = 4.00 \times 10^9 \text{ ergs/cm}^2/\text{s},$$

$$t_m = .040 \text{ s},$$

$$t_e = .075 \text{ s},$$

$$T_a = 300\text{K}.$$

The HULL code is used according to the sequence of steps outlined in Section III. Once the cell size DDX (cm) is chosen the time interval is determined either by the Courant condition or a condition on the quantity of energy added to cell 1, whichever is smaller. The energy added to the first cell in the time interval DDT (seconds) is restricted to be no greater than 10% of the energy already in the cell. This puts an upper bound on DDT; the Courant condition test for all cells results in another upper bound on DDT; the smallest of the two values is the one used for the current time-step.

A further feature of the calculation is an upper bound on the surface temperature  $T_1$ ; i.e., the cell-centered temperature in cell 1. If the cell size DDX is sufficiently small  $T_1$  can become greater than the vaporization temperature of the ground soil (2300K for this calculation). In that case a cut-off criterion is used; when  $T_1 > 2300\text{K}$  all of the flux

is absorbed in phase transition, and  $T_1$  is kept at 2300K.  $T_1$  remains at 2300K until the energy flux has decreased (become negative) sufficiently to cool cell 1 below 2300K.

A proper calculation for this effect would include two fluids and their energetics.

#### VIII. PRELIMINARY RESULTS

Figure 1 shows the results of a HULL calculation to determine the time required to reach 2300K in cell 1 for a range of cell lengths (1 to 30cm), using Eq.(6) for the energy flux. As expected,  $t \rightarrow 0$  for  $DDX \rightarrow 0$ ; on the other hand, there is a limiting value of  $DDX$  above which the temperature never reaches 2300K; in this case  $DDX = 33.31\text{cm}$ . The inflection point, at approximately  $t = .04\text{s}$ ,  $DDX = 19\text{cm}$ , is where  $d(DDX)/dt$  is a maximum; i.e., where changes in  $DDX$  result in a minimum change in  $t$ , and where both increases and decreases in  $DDX$  cause changes in  $t$  of equal magnitude. It is, perhaps, the proper choice for  $DDX$ . This is pure conjecture at this point, and may be worth pursuing in future work. The "correct" choice for  $DDX$  must be related to the physics associated with the deposition of energy in cell 1.

Figure 2 shows curves of temperature vs. distance at the moment when the temperature in cell 1 reaches 2300K; the parameter is cell length. It is seen that small  $DDX$  results in a rapid rise to 2300K with very little spread of the energy to the remainder of the gas. Larger  $DDX$  gives a slower rise to 2300K, with greater spreading of energy. For  $DDX > 33.1\text{cm}$  the succeeding values of  $DDX$  give curves that do not reach 2300K, and tend to follow the same locus as that for  $DDX = 33.1\text{cm}$ ; it is tempting to

conclude that this suggests one choose the maximum value of DDX as the "correct" cell length, at least for cell 1. Again, further work is needed to test this conjecture.

Figures 3 and 4 show temperature vs. distance for  $DDX = 20\text{cm}$  and a range of times. The region of temperature increase is confined to  $X < 100\text{cm}$ , attesting to the weak energy coupling between ground and air. Curves of pressure vs distance show small departures from atmospheric levels. The fluid velocity reaches values of  $1000\text{cm/s}$  within 10 to 20cm of the surface, and decreases to zero with distance from the surface. Both positive and negative values of velocity are computed, in accordance with the rise and fall of the energy flux, and the subsequent increase and decrease in density.

#### IX. RECOMMENDATIONS

This summer's effort represents a beginning toward building a thermal layer model that is both basic (minimally dependent on experimental correlation), and simple enough to be conveniently incorporated into the HULL code. Several items deserve further attention:

- 1) cell-size dependence becomes important when energy is added to cell 1 through the earth/air interface; this should be accounted for by including in the model some of the physics associated with the actual heat transfer process:
- 2) real fluid effects such as thermal conductivity, and viscosity may be important near the interface, where energy transfer is dominant:

- 3) the model should be extended to include energy exchange with any condensed phase (dust, water droplets, etc.) raised into the air by the fluid velocity:
- 4) ultimately, depart from the one-dimensional limitation to include two-dimensional flows; incorporate a no-slip boundary condition at the interface; this would be a major extension of HULL.

X. REFERENCES

1. F. A. Allahadadi, "Assessment of the State of the Art of Dust-Entrained Shock Physics," NMERI TAll-3, New Mexico Engineering Research Institute, University of New Mexico, Albuquerque, NM, August 1983.
2. J. T. Power, J. E. Mansfield, and R. T. Liner, "Precursor Sweep-up Dust Cloud Model and Thermal Layer Model Development," DNA 3876F, Defence Nuclear Agency, Washington, DC, 1975.
3. M. Rosenblatt, "Introduction to Nuclear Dust/Debris Cloud Formation," DNA 5832T, Defence Nuclear Agency, Washington, DC, July 1, 1981.
4. M. A. Fry, et al. "The HULL Hydrodynamics Code," AFWL-TR-76-183, Air Force Weapons Laboratory, Kirtland AFB, NM, Sept. 1976.
5. R. L. Bell, Major, and C. Westmoreland, "Elastic/Plastic HULL (EPHULL) Operation on the Cray Time Sharing System," AFWL-TR-83-6, April 1983.
6. C. E. Needham and J. E. Crepeau, "A Flux Dependent Thermal Layer Model," DNA 5538T, Oct. 31, 1980.
7. M. T. Sandford II, R. W. Whitaker, and R. C. Anderson, "Radiation Hydrodynamics of the Thermal Dust Layer," LA-UR-78-2411, Los Alamos Scientific Laboratory, Los Alamos, NM, 1978.
8. R. E. Durrett and D. A. Matuska, "The HULL Code: Finite Difference Solution to the Equations of Continuum Mechanics," AFATL-TR-78-125, Air Force Armament Laboratory, Eglin Air Force Base, Florida, Nov. 1978.

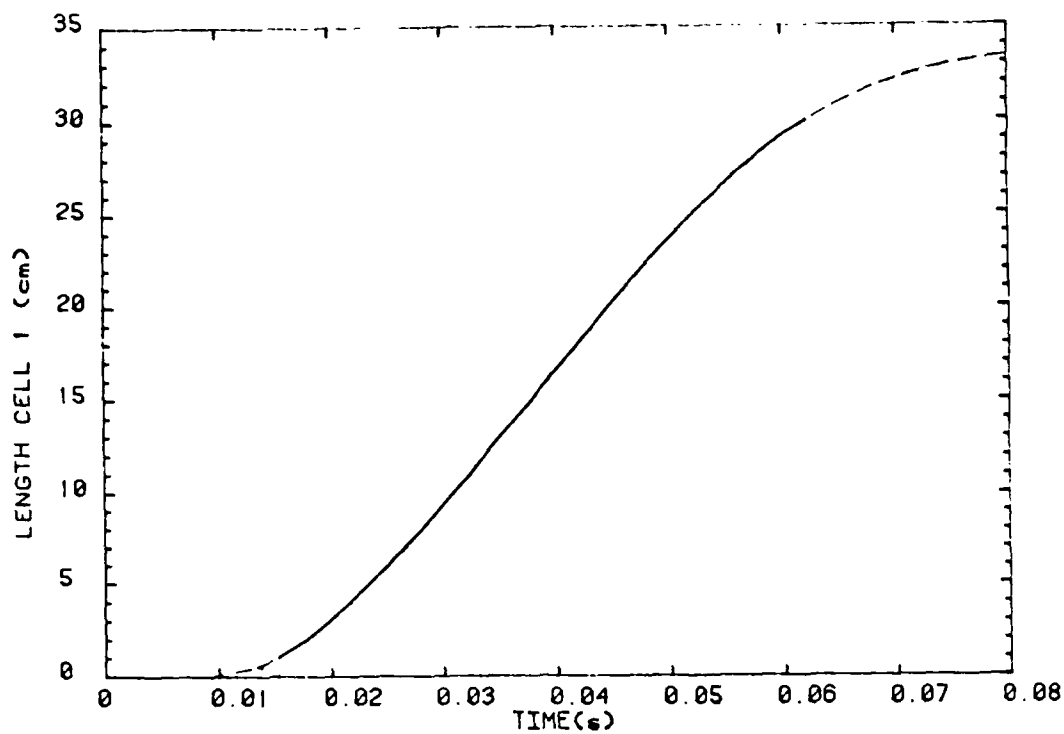


Fig. 1 TIME TO REACH 2300K IN CELL 1

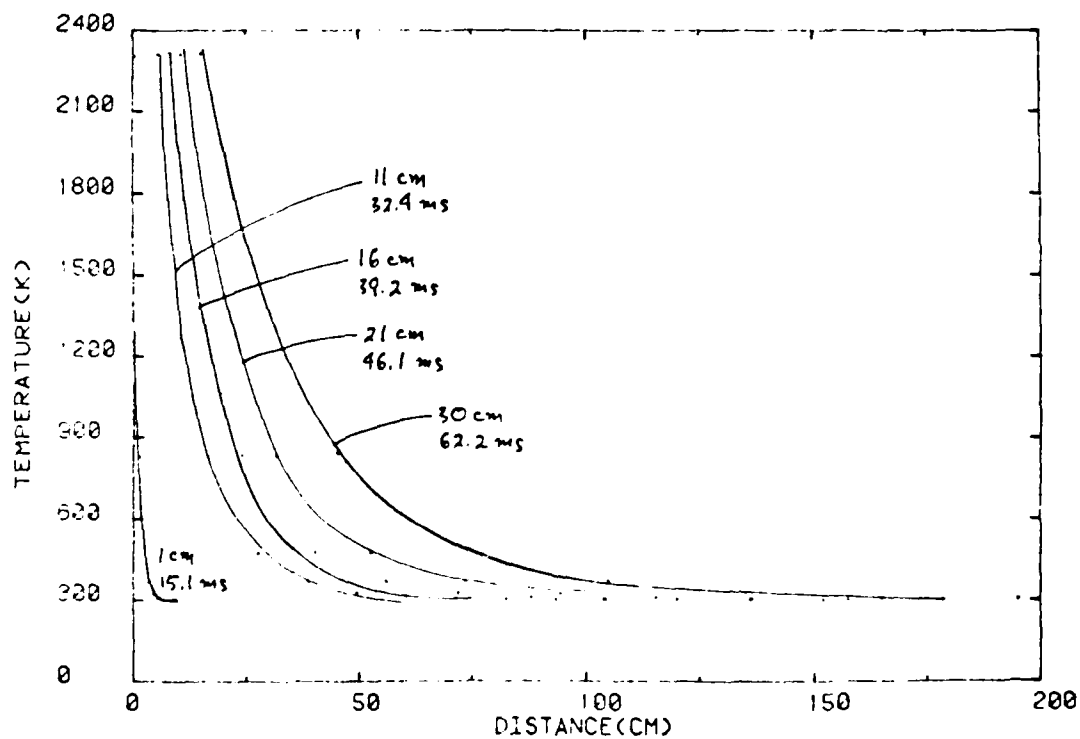


Fig. 2 TEMPERATURE VS DISTANCE FOR A RANGE OF UNIFORM CELL SIZES



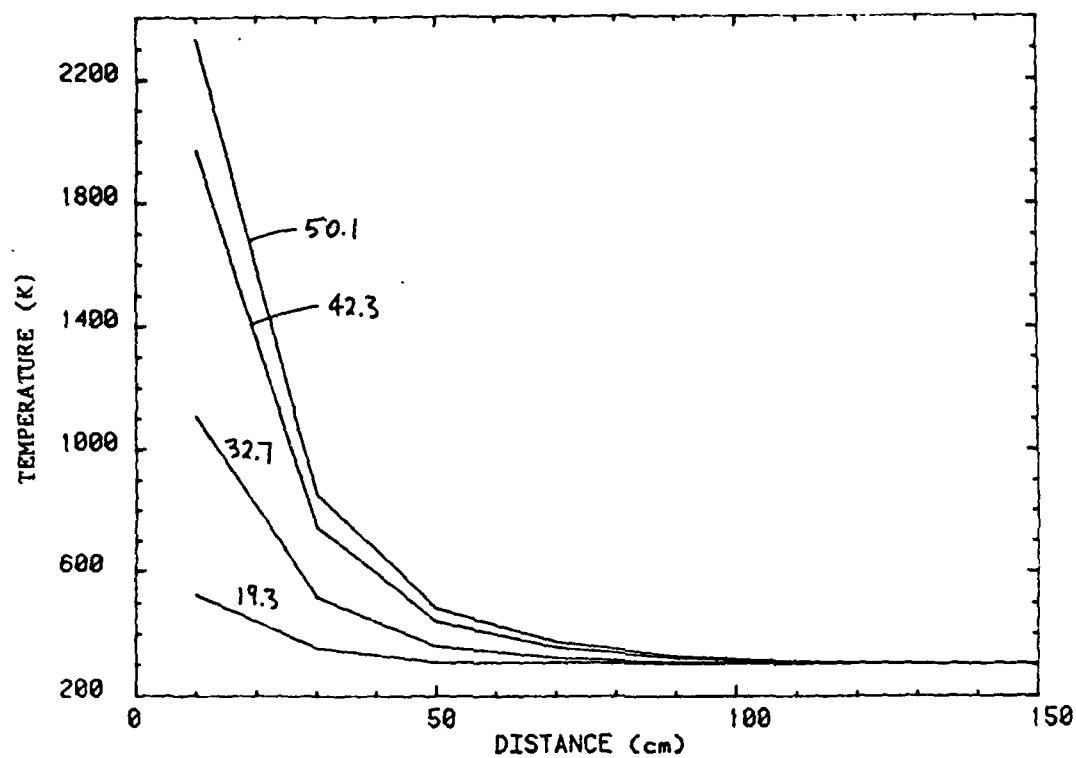


Fig.3 THERMAL LAYER EVOLUTION: 19.3 to 50.1 ms

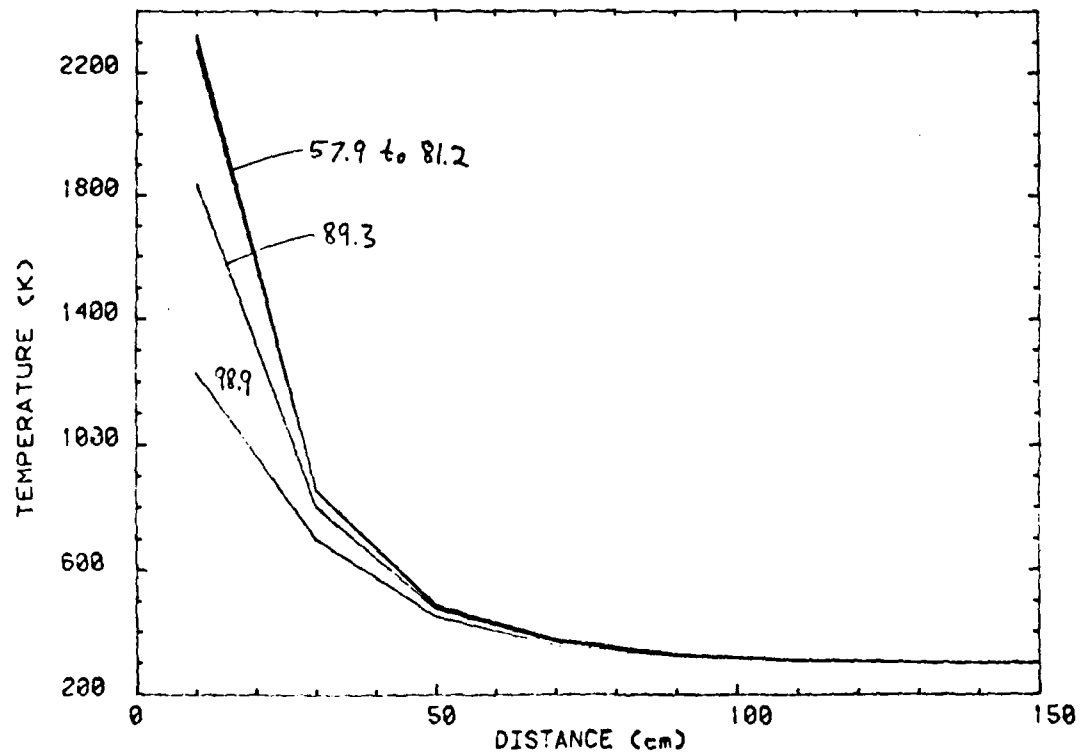


Fig.4 THERMAL LAYER EVOLUTION: 57.9 to 98.9 ms

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Computational Fluid Dynamics Grids-  
Flow Properties Interpretation Algorithm

Prepared by:	Dr. Madakasira G. Krishna
Academic Rank:	Professor
Department and University:	Department of Mathematics South Carolina State College
Research Location:	Arnold Engineering Development Center
USAF Research:	M. K. Kingery
Date:	August 10, 1984
Contract No:	F49620-82-C-0035

Computational Fluid Dynamics Grids-  
Flow Properties Interpretation Algorithm

by  
Madakasira G. Krishna

ABSTRACT

Plane of intersection between a specified (arbitrary) plane and an arbitrary three-dimensional grid on which internal flow is simulated is developed. Flow properties of interest are interpolated at a network of points on the intersected surface. A computer code is developed which needs very little user interference. Suggestions for further research in this area are offered.

## I INTRODUCTION:

Designers of exhaust nozzles and other complex parts need to know pressure, heat transfer and shear distributions to support design activities. Traditionally these are obtained by experimental data. Significant reduction in design costs can be obtained if calculated results based on finite difference procedures are used to verify and supplement experimental results.

Many computational fluid dynamics problems associated with internal flow simulation involve complex three-dimensional grids (1). The geometries of the external surfaces and internal obstacles involved in these problems are complex. Visualization of grid lines is not straight forward in the case of these three-dimensional grids.

When we consider internal flow in which there is a difference in plane expansion (as in the case of a vectored exhaust nozzle (Figure 1)), it is useful to know whether viscous effects should be taken into consideration or not at a given section. This knowledge will help in understanding whether to use a Euler solver or a much more expensive Navier-Stokes solver in a given region.

## II OBJECTIVES:

Compute a surface of intersection between a specified (arbitrary) surface and an arbitrary three-dimensional grid on which flow is being simulated. Evaluate a network of points on the intersected surface and interpolate flow properties of interest at these points. In this project instead of an arbitrary surface a plane section will be considered. In order to test the algorithm a simple three-dimensional grid of a cylinder (Figure 2) is considered. But the algorithm can be used for finding a plane section of any grid which does not involve interior obstacles. Section of a plane with grids involving interior obstacles is left for later investigation.

## III SECTION OF A PLANE WITH A THREE-DIMENSIONAL GRID:

Conceptually finding a section of an arbitrary plane with grid appears simple. Logically it is a hard problem because the plane does not pass through the grid points only and the mathematical description of the geometry of the solid is not known.

Given three arbitrary points the equation of plane

$$ax + by + cz + d = 0 \quad \dots(1)$$

is determined by solving for a, b, c, and d. The code is written such that the user can find out at the outset whether the plane under consideration intersects the user supplied grid or not.

When the plane cuts the grid three possible situations arise.

1. A grid point may be on the plane or very near the plane.

2. The plane passes through a grid volume.
3. A grid point is too far from the plane

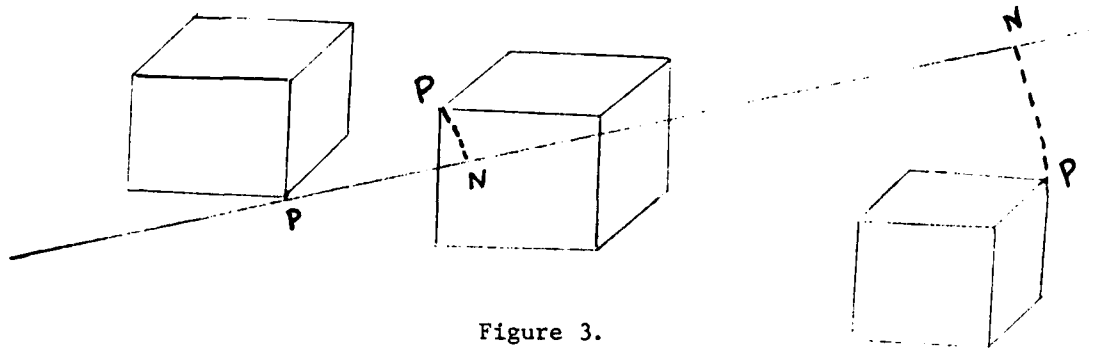


Figure 3.

If a point  $(x_1, y_1, z_1)$  on the grid satisfies the condition:

$$|ax_1 + by_1 + cz_1 + d| \leq 0.001 \quad \dots(2)$$

then  $(x_1, y_1, z_1)$  is considered to be a point on the section. When

$P(x_1, y_1, z_1)$  is not on the plane we find the length of perpendicular PN from P to the plane

$$PN = \frac{|ax_1 + by_1 + cz_1 + d|}{\sqrt{a^2 + b^2 + c^2}} \quad \dots(3)$$

and compute the distances

$$\begin{aligned} Dx &= |x(I+1, J, K) - x(I, J, K)| \\ Dy &= |y(I, J+1, K) - y(I, J, K)| \\ Dz &= |z(I, J, K+1) - z(I, J, K)| \end{aligned} \quad \dots(4)$$

and find

$$D_{xx} = \text{MINIMUM } (Dx, Dy, Dz) \quad \dots(5)$$

If PN is greater than Dxx we discard  $(x_1, y_1, z_1)$ . This will eliminate all points on the grid which are not near the plane.

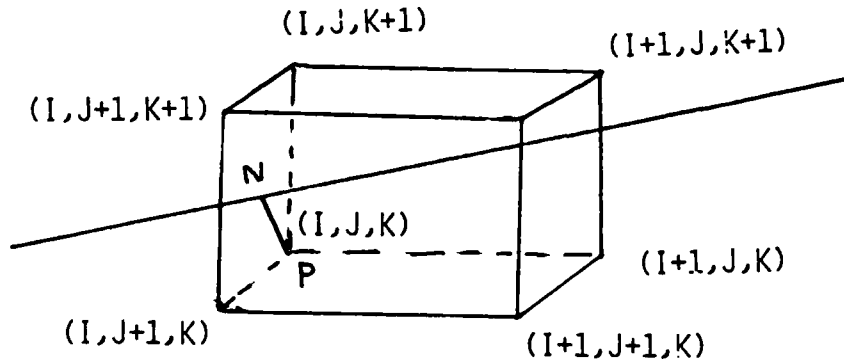


Figure 4.

When  $PN < D_{xx}$  the plane passes through a grid volume. In this case we drop a perpendicular PN from  $P(x_1, y_1, z_1)$  to the plane. We find the coordinates of the foot of the perpendicular  $N(\bar{x}, \bar{y}, \bar{z})$  using the equation of line PN given by

$$\frac{\bar{x}-x}{a} = \frac{\bar{y}-y}{b} = \frac{\bar{z}-z}{c} = \lambda \quad \dots(6)$$

and the fact that  $(\bar{x}, \bar{y}, \bar{z})$  is on the plane (1) we find  $\lambda$  as

$$\lambda = - \frac{(ax_1 + by_1 + cz_1 + d)}{a^2 + b^2 + c^2} \quad \dots(7)$$

and hence

$$\begin{aligned} \bar{x} &= x_1 + a\lambda \\ \bar{y} &= y_1 + b\lambda \\ \bar{z} &= z_1 + c\lambda \end{aligned} \quad \dots(8)$$

Using this method a collection of points on the plane can be obtained which represents the section of the plane with the grid. It is hard to determine the boundary of this section because it is difficult to keep track of the running variables I, J, K which vary from point to point.

In order to overcome this difficulty a different method is used in which we take advantage of the method used in generating a three-dimensional grid. In generating a three-dimensional grid six surfaces of the solid are considered. (1) On each one of these surface I or J or K has a constant value as shown in Figure (5)

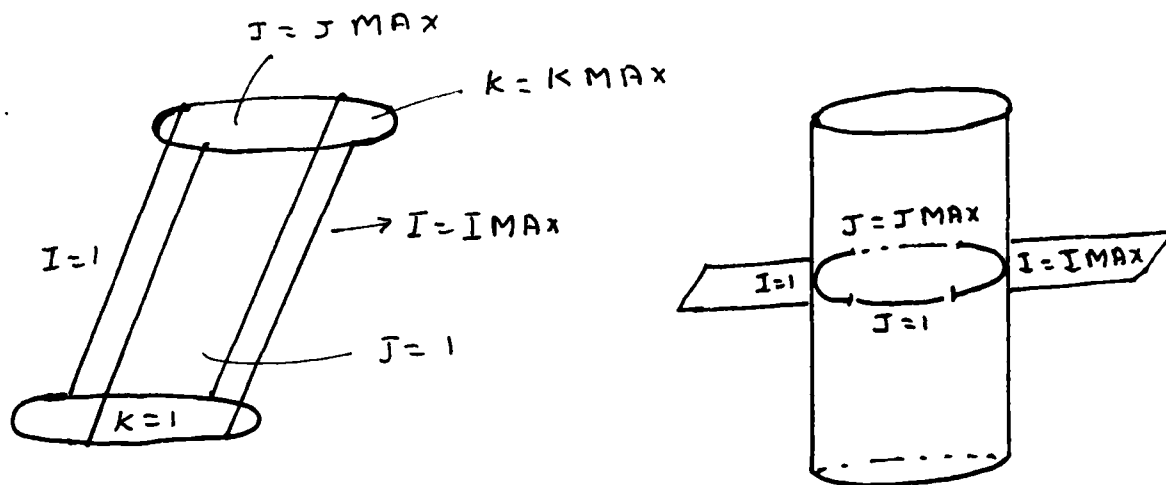


Figure 5.

First we assume that the plane cuts the grid along  $I = 1$ ,  $I = I_{MAX}$ ,  $J = 1$  and  $J = J_{MAX}$ . When  $I = 1$  we have to keep track of only two running variables  $J$  and  $K$  which is not difficult to do. The  $I = 1$  boundary is obtained by collecting all points satisfying

$$|ax + by + cz + d| \leq 0.001 \quad \dots (9)$$

and by considering the foot of the perpendicular  $N(\bar{x}, \bar{y}, \bar{z})$  from all points  $P(x_1, y_1, z_1)$  on the grid which satisfy

$$PN = \frac{ax + by + cz + d}{\sqrt{a^2 + b^2 + c^2}} \leq D_{xx} \quad \dots (10)$$

Here  $\bar{x}$ ,  $\bar{y}$ ,  $\bar{z}$  are given by Equation (8) and  $D_{xx}$  by Equation (5).



Similarly we determine points on the plane section corresponding to  $I = I_{MAX}$ ,  $J = 1$  and  $J = J_{MAX}$ . Figures 6, 7, 8 and 9 shows the boundaries corresponding to  $I = 1$ ,  $I = I_{MAX}$ ,  $J = 1$  and  $J = J_{MAX}$  for the intersection of plane  $3x + 2y - 20z + 100 = 0$  with a cylindrical grid. These boundaries have 25 to 28 points on them and they are not very smooth. If the end points of the four boundaries do not match we take their mean and come up with a closed boundary. We choose some 20 points on the  $I = \text{Constant}$  Boundary and 15 points on  $J = \text{Constant}$  Boundary and fit a smooth curve (2) to pass through them. Using the points on the four boundaries we come up with a network of points in the section (Figure 10) using a subroutine GRDSD (1).

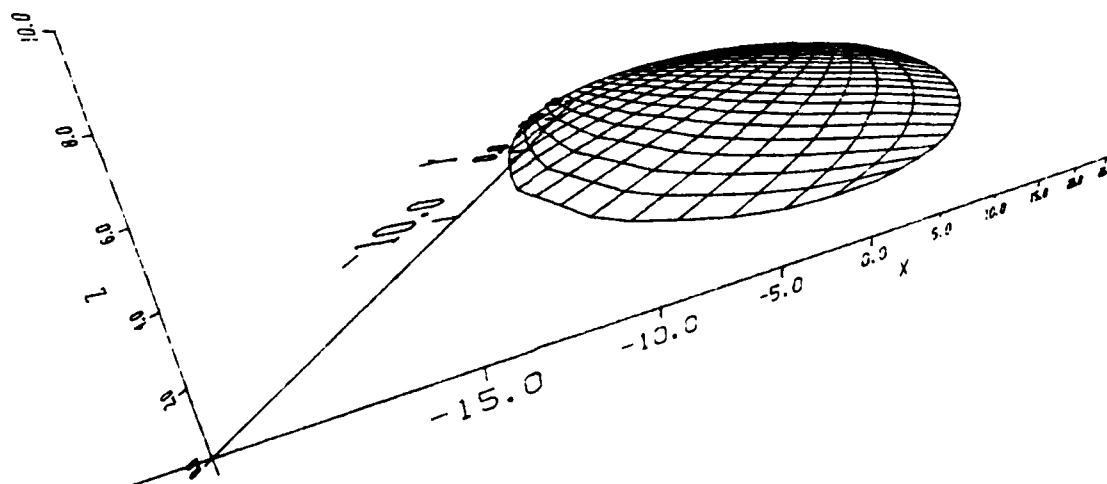


Figure 10.

SECTION OF PLANE  $3x + 2y - 20z + 100 = 0$  WITH  
CYLINDER ALONG WITH NETWORK OF POINTS

The next step is to interpolate the flow properties at these network of points. As before we consider all points  $(x_i, y_i, z_i)$  on the grid satisfying Equation (9) and relation (10) and their distances  $d_i$  from a specific point  $L(x_1, y_1, z_1)$  on the network  $d_i$  is given by

$$d_i = \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2 + (z_i - z_1)^2} \quad \dots(11)$$

To interpolate we use only those points  $(x_i, y_i, z_i)$  whose distance  $d_i$  from  $L(x_1, y_1, z_1)$  satisfy the condition

$$d_i \leq D_{xx} \quad \dots(12)$$

Here  $D_{xx}$  is given by equation (5). In order to give more weight for points which are close to the point  $L$ , we compute

$$\begin{aligned} D_1 &= 1 - \frac{d_1}{\sum d_i} \\ D_2 &= 1 - \frac{d_2}{\sum d_i} \quad \text{etc} \quad \dots(13) \end{aligned}$$

and write

$$U(x_1, y_1, z_1) = \frac{\sum U_i D_i}{\sum D_i} \quad \dots(14)$$

where  $U_i$  are velocity components at  $(x_i, y_i, z_i)$  on the grid in the  $x$  direction

The other flow parameters can be found similarly.

The other two options

$$J = 1, J = J_{MAX}$$

$$K = 1, K = K_{MAX}$$

and

$$I = 1, I = I_{MAX}$$

$$K = 1, K = K_{MAX}$$

are handled in the same way. The code is written in such a way that the user can find out as soon as the plane is defined which one of the three options to use. Even though a simple three-dimensional grid is used to test the algorithm, the code can be used with more complex grids. The code

can also be used to find a section of a surface with a grid as long as the surface is defined analytically or numerically.

#### IV CONCLUDING REMARKS:

A computer code is written to find any section of a plane or surface with a three-dimensional grid. The code has three different options. The user can select one of these three options depending on where the defined plane intersects the grid. Sections of a cylindrical grid with two planes are presented in Figures 10 and 11. By comparing the flow parameter at a particular section one can decide whether to use an Euler solver or a more expensive Navier-Stokes solver in grids where there is plane expansion.

#### RECOMMENDATIONS:

Design costs of complex parts such as vectored exhaust nozzles can be reduced significantly if available experimental results can be supplemented by finite difference procedures. In using finite difference procedures if one knows whether viscous effects should be taken into consideration or not in a particular section of the grid can save much expense. This knowledge is gained by comparing the flow parameters at a given section. In this project three-dimensional grids which do not involve internal obstacles are considered. But in practice many internal flow problems involve internal obstacles. I propose to write a code to find a plane section of a three-dimensional grid in which internal obstacles are involved.

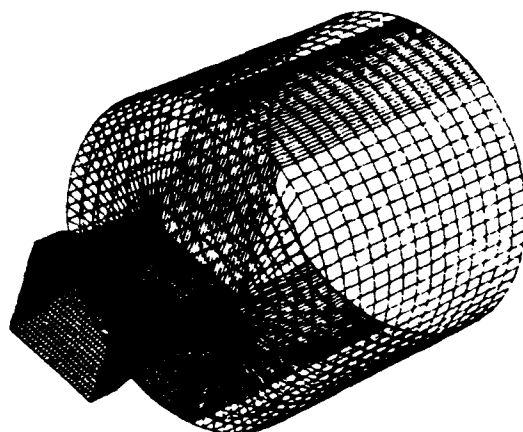


Figure 1.

PARTIAL THREE-DIMENSIONAL GRID OF A VECTORED EXHAUST NOZZLE

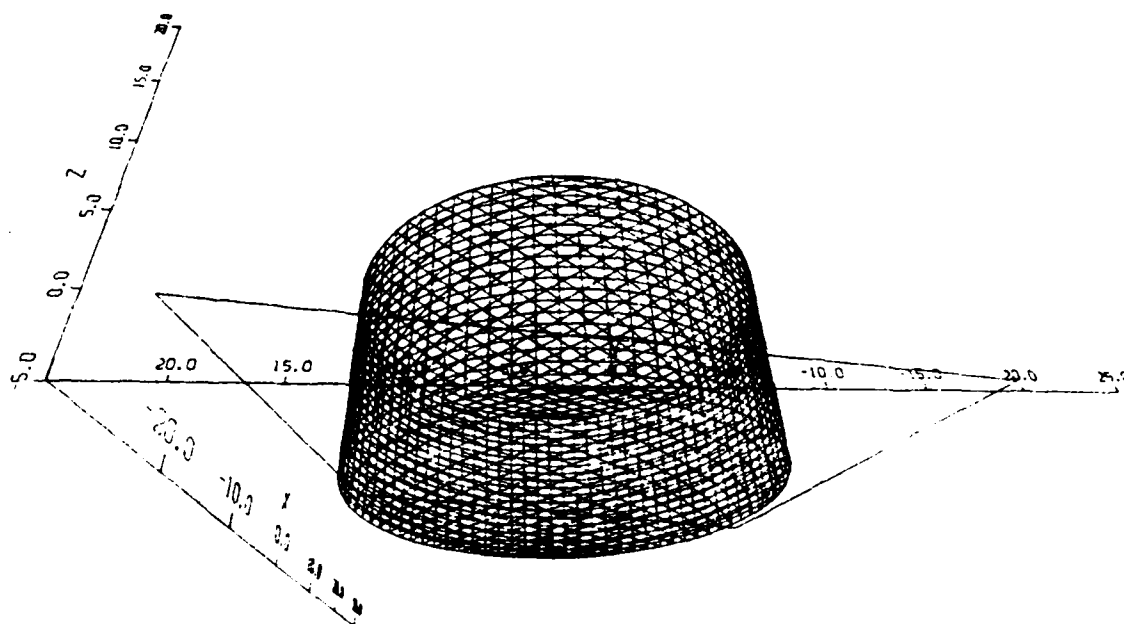


Figure 2.

THREE-DIMENSIONAL GRID OF A CYLINDER ALONG WITH A PLANE THAT INTERSECTS

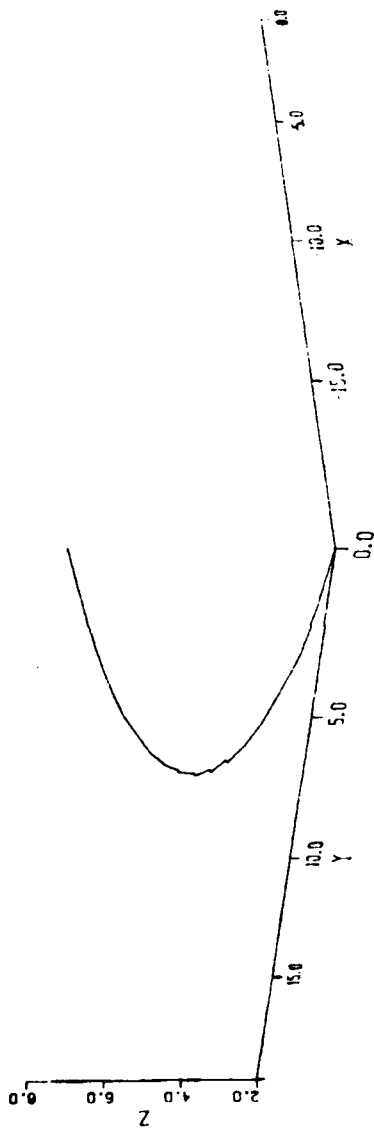


Figure 6.

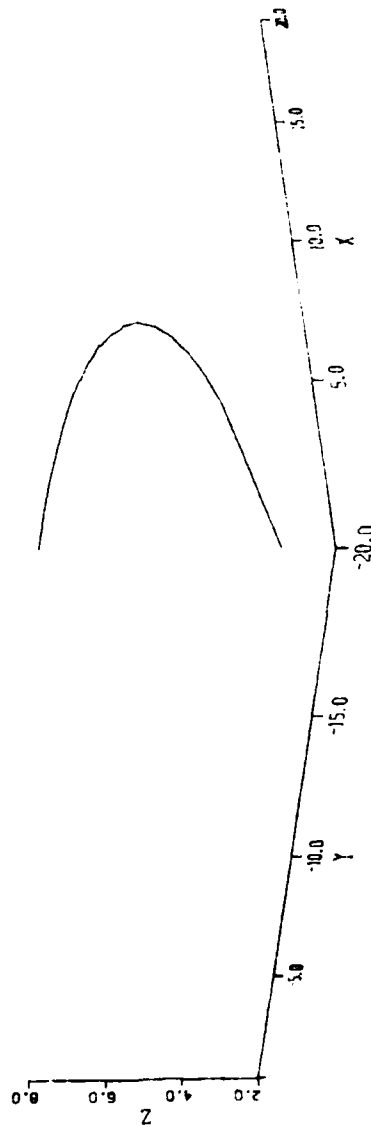


Figure 7.

SECTION OF PLANE  $3x + 2y - 20z + 100 = 0$  WITH GRID  
FOR CYLINDER WHEN  $I = 1$  AND  $I = \text{IMAX}$

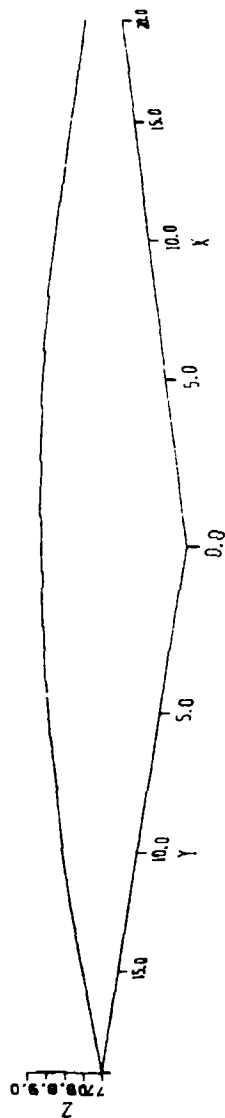


Figure 8.

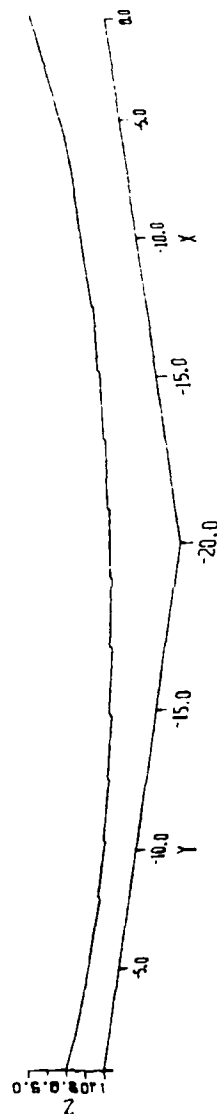


Figure 9.

SECTION OF PLANE  $3x + 2y - 20z + 100 = 0$  WITH GRID  
FOR CYLINDER WHEN  $J = 1$  AND  $J = J_{MAX}$

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

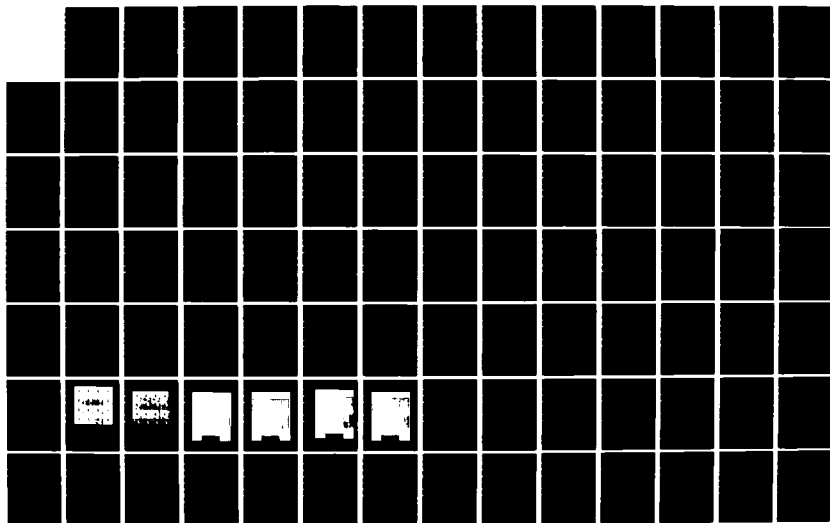
9/13

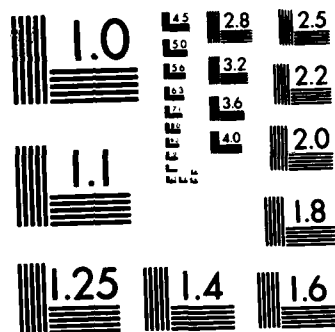
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A



# THE OTHER 2 OPTIONS

$J = 1, J = JMAX$

$K = 1, K = KMAX$

$I = 1, I = IMAX$

$K = 1, K = KMAX$

ARE HANDLED IN THE SAME WAY

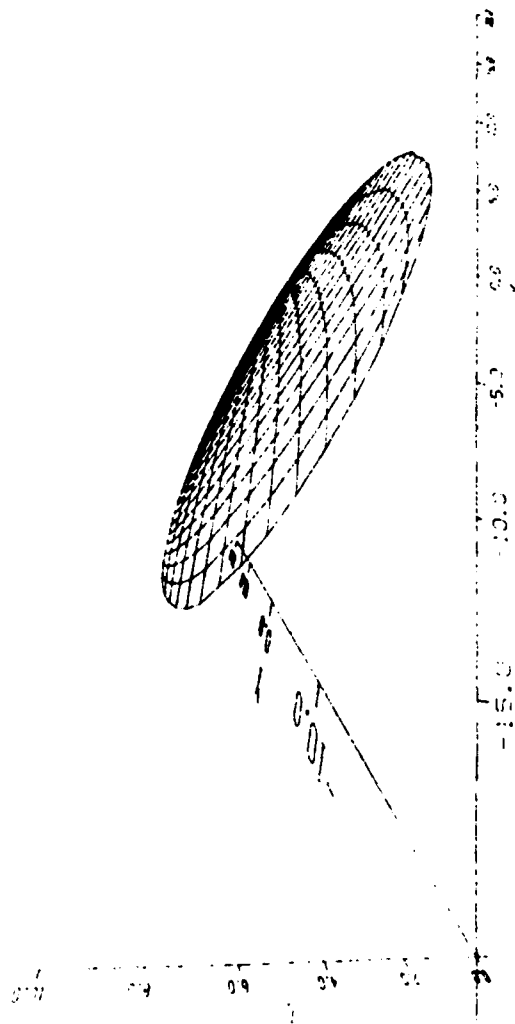


Figure 11.

SECTION OF PLANE  $2x - y + 10z - 50 = 0$

WITH CYLINDRICAL GRID

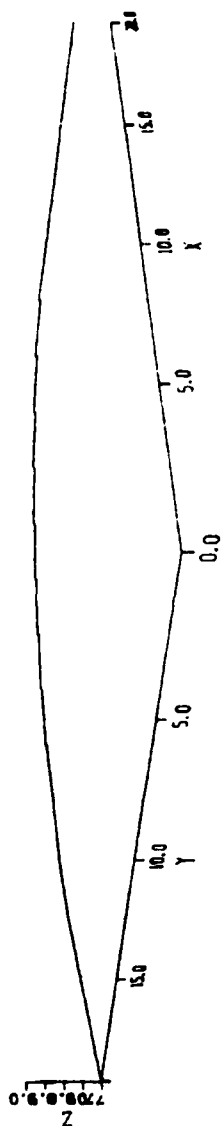


Figure 8.

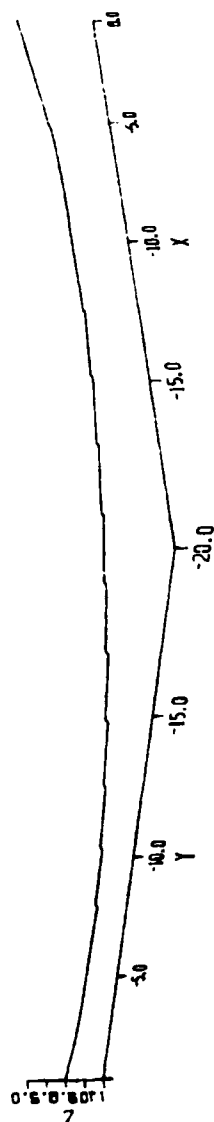


Figure 9.

SECTION OF PLANE  $3x + 2y - 20z + 100 = 0$  WITH GRID  
FOR CYLINDER WHEN  $J = 1$  AND  $J = J_{MAX}$

#### ACKNOWLEDGEMENT

The Author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Arnold Engineering Development Center, Arnold Air Force Station, Tennessee. He would like to acknowledge the Center, in particular the Computational Division of Sverdrup Technology, Inc., for its hospitality and excellent working conditions.

Finally he would like to thank Dr. Bharat Soni for suggesting this area of research and for his collaboration and guidance, and he would like to acknowledge helpful discussions with Wilbur Armstrong.

#### REFERENCES

1. Soni B. K., "Users Guide For Internal Flow Grid Generation Applications"  
Users Guide, Sverdrup Technology, Inc.
2. Akima, H., "A New Method of Interpolation And Smooth Curve Fitting  
Based On Local Procedures" JACM, 17(4) 1970, pp 589-602.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

DEVELOPMENT OF AN OPTICAL MULTICHANNEL ANALYZER SYSTEM

Prepared by:	Dr. Raj. M. Krishnan
Academic Rank:	Associate Professor
Department and University:	Department of Physics North Texas State University
Research Location: Beam	School of Aerospace Medicine, Brooks AFB
USAF Research	Dr. John Taboada
Date:	September 12, 1984
Contract No:	F49620-82-C-0035

DEVELOPMENT OF AN OPTICAL  
MULTICHANNEL ANALYZER SYSTEM

by

Raj. M. Krishnan

ABSTRACT

The objective of the research is to develop an optical multi-channel analyzer system which is capable of recording Raman Scattered light over a wide band width simultaneously, so that time dependent phenomena can be studied. A varian triple monochromator was interfaced to a Hamamatsu Camera through an image intensifier. The output of the camera was interfaced to a PDP 11/23 plus computer. The varian spectrometer was modified so that both the multichannel capability and the traditional photon counting system could be used. These objectives were achieved. Suggestions for improving the system are offered.

### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the School of Aerospace Medicine, Brooks AFB, Texas.

I would like to thank Dr. John Taboada for suggesting this area of research and for his collaboration and guidance.

## I. INTRODUCTION:

The transmission of energy in biological systems has been a puzzle, and it has attracted considerable attention during the past decade. The main issue centered around the fact that energy transmission via excited molecular vibrations is presumed to have too short a life time. In 1977 Davydov<sup>1</sup> proposed a potential solution to this problem. Davydov suggested that nonlinear interatomic interactions can produce solitary waves, called "solitons". Solitons can transport energy over large distances without appreciable loss of energy. Solitons were first observed as fast moving water loops in narrow channels filled with water. They traveled over long distances without any diminishing of amplitude.

In alpha-helical proteins the peptide groups ( $\text{H}-\text{N}-\text{C}=\text{O}$ ) form three molecular chains. The peptide groups are held together by hydrogen bonding between H and O, providing a rigid structure for the  $\alpha$ -helix proteins. Davydov showed that when nonlinear interactions between atoms are taken into account solitary waves can be formed. These solitary waves or "solitons" have long life span and can transfer energy with high efficiency.

Davydov solved the problem by assuming that the interaction between the chains is weak. Thus, he reduced the problem to a single chain. He showed that if the nonlinear interaction exceeds a certain strength, "soliton" formation is possible.



Hyman, McLaughlin and Scott<sup>2</sup> extended the Davydov model and made a detailed numerical calculation and showed that the nonlinearity is large enough to lead to the formation of solitons in  $\alpha$ -helix proteins.

We shall briefly discuss the soliton equations. The atomic structure of  $\alpha$ -helix protein follows the sequence: etc., -N-C-C-N-C-C-N- etc. with a pitch of 5.4 Å. Superimposed on this basic structure are three "spines" which are almost longitudinal with the sequence:

etc. ---N-C = O --- N-C = O ---N-C=O----etc.

The following equations describe the coupled propagation of bond energy probability density,  $A_{n\alpha}$ , and longitudinal displacement of an amide-I bond,  $B_{n\alpha}$

$$i \frac{dA_{n\alpha}}{d\tau} + J(A_{n+1,\alpha} - 2A_{n\alpha} + A_{n-1,\alpha}) - k_1(B_{n+1,\alpha} - B_{n-1,\alpha})A_{n\alpha} = L E + \{ k_2 [B_{n-1,\alpha} A_{n-1,\alpha} - B_{n+1,\alpha} A_{n+1,\alpha} + B_{n\alpha} (A_{n+1,\alpha} - A_{n-1,\alpha})] - 2J A_{n\alpha} + N F_N + P F_P + Q F_Q + R F_R + S F_S + T F_T + U F_U + V F_V + X F_X + Z F_Z \}$$

$$\frac{d^2 B_{n\alpha}}{d\tau^2} - (B_{n+1,\alpha} - 2B_{n\alpha} + B_{n-1,\alpha}) =$$

$$k_3 (|A_{n+1,\alpha}|^2 - |A_{n-1,\alpha}|^2) + \{k_4 [A_{n\alpha}^* (A_{n+1,\alpha} - A_{n-1,\alpha}) + A_{n\alpha} (A_{n+1,\alpha}^* - A_{n-1,\alpha}^*)]\}.$$

In these equations ( $J, L, \dots, Z$ ) are dipole-dipole interaction energies between pairs of amide-I bonds. The subscript  $n$  represents unit cells along the helix while  $\alpha$  specifies one of the three spines. The constants  $k_1 \dots k_4$  represent nonlinear coupling between bond energy dispersion and longitudinal sound.  $\tau$  is the time variable. The numerical solution to these equations is given in Reference 2.

The question now arises as to how one could detect the soliton excitations experimentally. During the years 1975 to 1979 Webb and his co-workers<sup>3</sup> attempted to obtain the Laser-Raman Scattering from living cells. A large number of lines were observed up to wave number  $3000 \text{ cm}^{-1}$ . However these were only observed from metabolically active cells. In particular there were several lines seen below  $200 \text{ cm}^{-1}$  (Fig. 1). The extension of the Davydov model, without any free parameters, by Scott<sup>4</sup> also predicted soliton excitations which should be observed below  $200 \text{ cm}^{-1}$  in Raman Scattering. A comparison of experiment and theory is shown in Table 1. The agreement is impressive.

Even though the above discussions point to tentative verification of Davydov solitons, the nature of the problem involving living cells suggest that additional experimental evidence is necessary.

Recently Taboada and Mrotek<sup>5</sup> found new experimental evidence for solitons. They observed Raman Scattering from HEP-2, a human nasal and y-1, a mouse adrenal, cancer cells. The results support the Davydov soliton theory.

In spite of the successes noted above, one thing that is consistently seen in the experiments is the time dependency of the Raman activity, when one deals with living cells.

In the conventional Raman Spectrometers using photon counting techniques one does not observe the whole spectrum simultaneously. Since a scanning method is used, if the sample changes as a function of time one cannot make definite conclusions. The method used in the past is to repeat the experiment many times and analyze the results using histogram techniques.

The situation can be remedied by using optical multichannel systems. In these systems one opens up the spectrometer slit as much as possible and the spectrum is recorded using camera or photodiode arrays. The procedure is made more attractive by the fact that modern computers can acquire the data extremely fast. The project is concerned with modifying a varian triple monochrometer, so that a video-camera can record the data simultaneously over a wide band.

## II. OBJECTIVES

The main objective of this project is to modify a varian triple monochromator to include an optical multichannel analyzing capability and a scanning-photon counting capability. Our specific objectives were

- 1) To install a videocamera at the output port of the spectrometer through an image intensifier.
- 2) To interface the camera to a PDP 11/23 plus computer and develop the appropriate software.
- 3) To install a photon counting system which can be interfaced to the PDP 11/23 plus computer.

## III. OPTICAL MULTICHANNEL ANALYZER (OMA)

The primary advantage of the OMA is the ability to instantaneously capture a complete optical spectrum. A schematic of the system is shown in Fig. 2. A spectrophysics Argon-ion laser is used as the light source. The scattered light is analyzed through a Varian triple monochromator. The slit can be opened up to  $50 \text{ cm}^{-1}$ . The scattered light is intensified by a Varo image intensifier. The intensified light is focused onto a Hamamatsu C-1000 camera. The camera, with the camera control, is capable of digitizing the image up to  $1024 \times 1024$  resolution. The system was successfully developed and scattered light from Acetone and water was observed. We also took data. It was found that the edges of the slit introduced a profile on the spectrum.

#### IV. SOFTWARE FOR THE OMA

The manufacturer of the camera provided a demonstration software system to be used in the RT11 operating system. The software had to be modified to fit our application. In particular fast acquisition of the data was one of the main objectives. The driver was written in Macro-11 assembly language. The main program is called VICTRL ( , , , , , ) which has seven arguments. The first variable defines the function to be performed. This program can be called from a FORTRAN program. The time consumed in acquiring data is a function of the resolution and the number of integrations to be performed. The number of integrations is determined by the intensity of the scattered light.

#### V. PHOTON COUNTING SYSTEM

The Schematic for the photon counting system is shown in Fig. 3. Test runs were made and the results are shown in Fig. 4. The system is found to be quite sensitive as can be seen from Fig. 4, where the water band  $1650 \text{ cm}^{-1}$  is clearly recorded.

#### VI. RECOMMENDATIONS

It was found that the number of integrations needed to obtain sufficient data was large. This made data acquisition too slow. It was decided to use two image intensifiers in tandem, thus increasing the intensity by a factor of 1000 - 10,000.

The edge effect must be dealt with by applying a correction to the data. This must be included in the software. One can perhaps observe the pattern by illuminating the slit with monochromatic light and correct for the edge effect using this data.

In order to use the system most efficiently it will be advantageous to increase the slit width, so that scattered light over several hundred wave numbers can be recorded at the same time.

TABLE I

Line	Experiment $\text{cm}^{-1}$	Theory	
		Excitation	$\text{cm}^{-1}$
1	45	$2E_1$	34
2	52	$3E_1$	51
3	63	$4E_1$	68
4	85	$5E_1$	85
5	90	$E_2 - 2E_1$	91
6	108	$E_2 - E_1$	108
7	123	$E_2$	125
8	152	$E_2 + 2E_1$	159
9	182	$E_2 + 3E_1$	176

Lines observed in the spectrum of *E. Coli* (Reference 3) are compared to the theoretical values calculated by Scott (Reference 4) using the Davydov Soliton concept.  $E_1$  and  $E_2$  represent the two excitations.

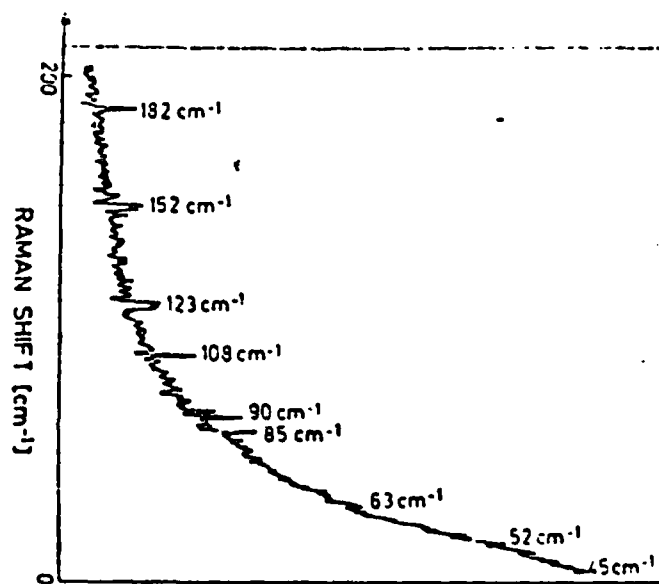


Fig. 1. Laser-Raman Spectrum from metabolically active E. Coli (Reference 3)



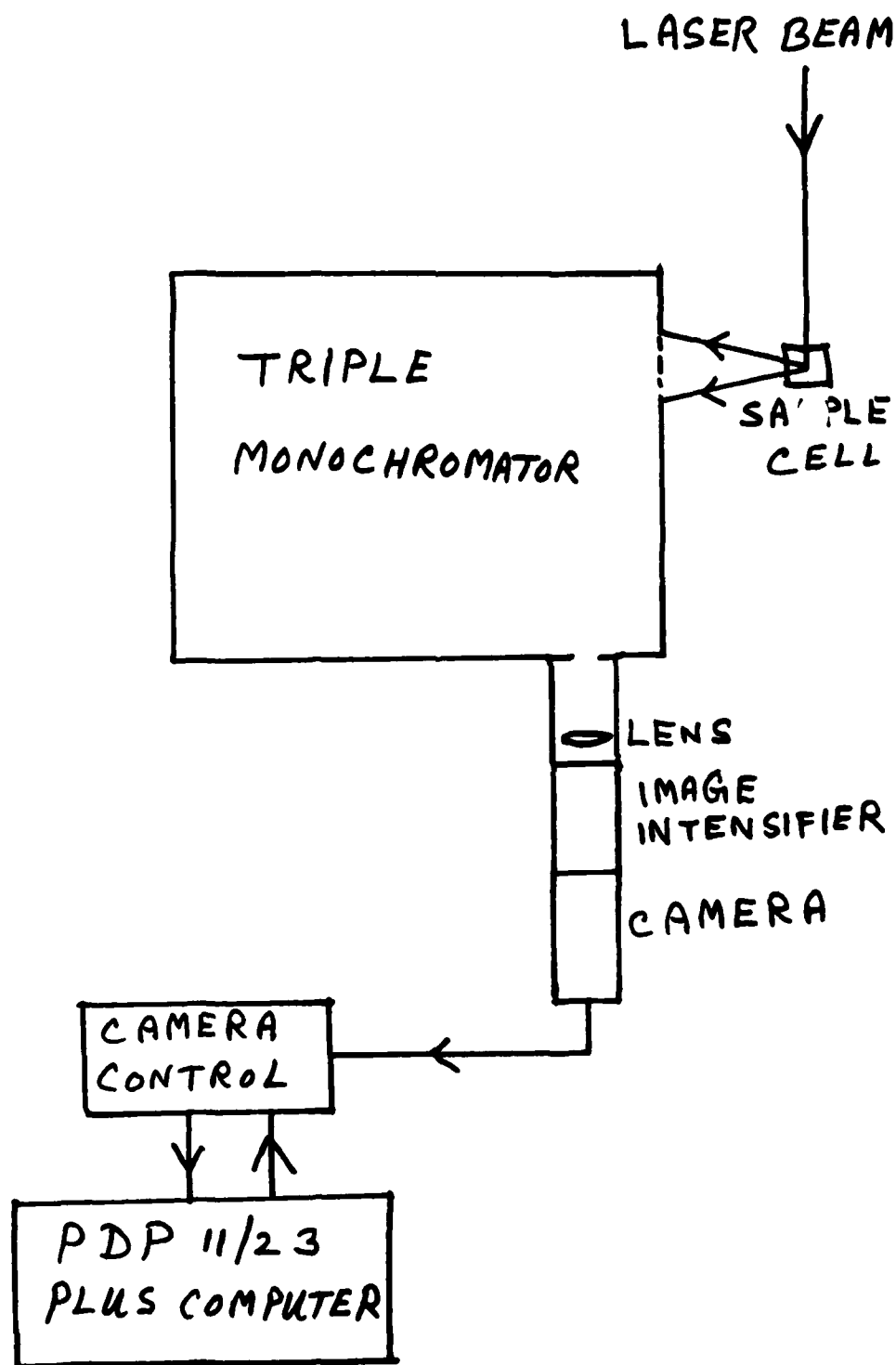


Fig.2. Schematic of the Optical Multichannel Analyzer.

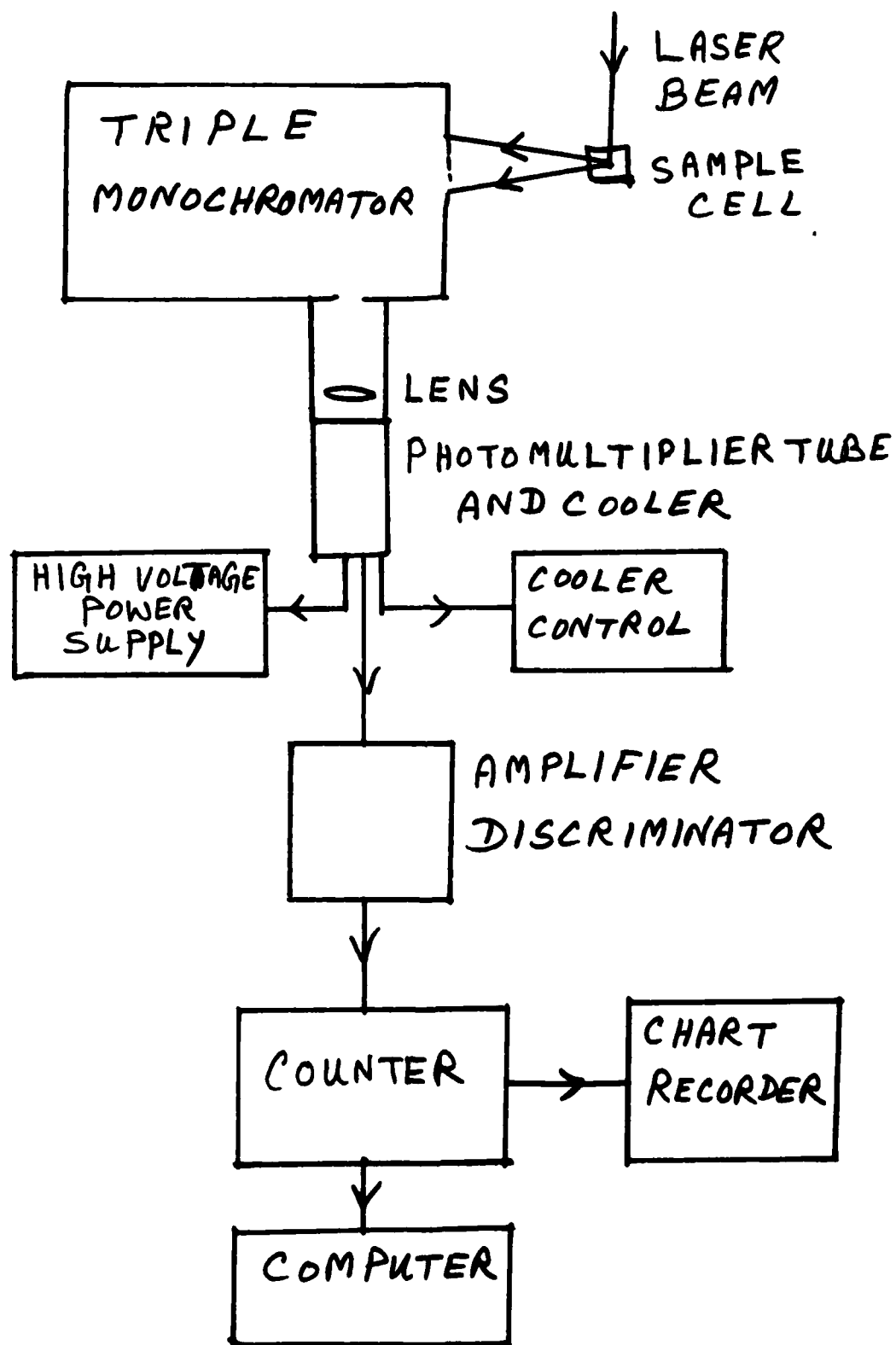
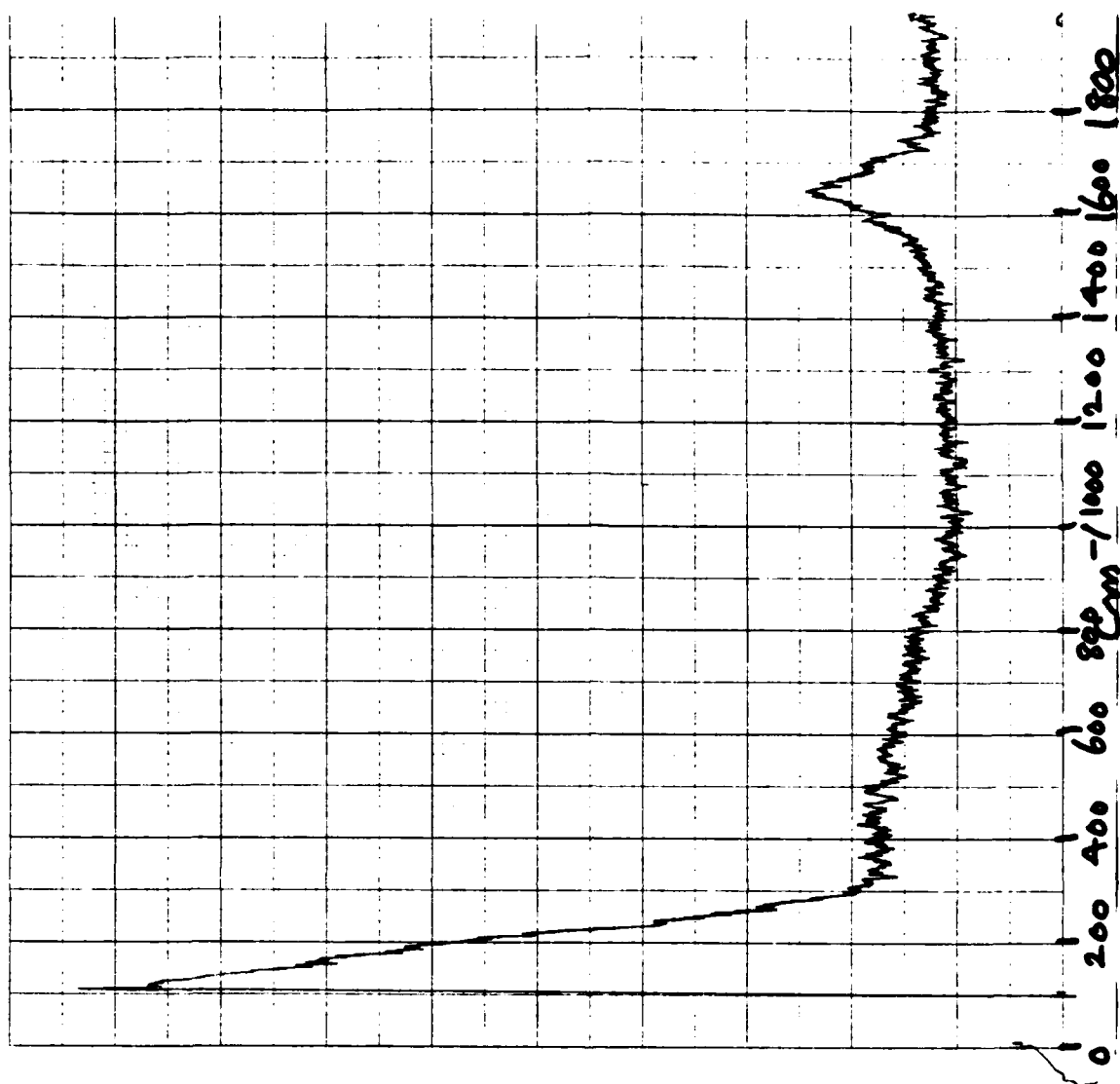


Fig. 3. Schematic of the Photon Counting System.



NY 1101-SP5

Fig. 4. Raman Spectrum of Water.

#### REFERENCES

1. A. S. Davydov, "Solitons and Energy Transfer along Protein Molecules", J. Theor. Biol., Vol. 66, pp 379-387, 1977.
2. J. M. Hyman, D. W. McLaughlin and A. C. Scott, "On Davydov's Alpha-Helix Solitons", Physica 3D, pp 23-44, 1981.
3. Sydney J. Webb, "Laser Raman Spectroscopy of Living Cells", Physics Reports, Vol. 60, pp 201-224, 1980.
4. A. C. Scott, "The Laser Raman Spectrum of a Davydov Soliton", Physics Letters, Vol. 86A, pp 60-62.
5. J. Taboada and J. Mrotek, "Evidence for Davydov Solitons in the Anti-Stokes Raman Spectra of Living HEP-2 and Y-1 Cancer Cells", Bull. Am. Phys. Soc., Vol. 29, pp 930, 1984.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM  
Sponsored by the  
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH  
Conducted by the  
SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

INTERLAMINAR SHEAR TESTING  
OF  
CARBON/CARBON COMPOSITES

Prepared by:	William Kyros, Ph.D.
Academic Rank:	Associate Professor
Department and University:	Mechanical Engineering Department University of Lowell
Research Location:	Materials Laboratory, AFWAL/MLBC Wright-Patterson Air Force Base, Ohio 45433
USAF Research:	Mr. Glenn Ormbrek
Date:	September 28, 1984
Contract Number:	F49620-82-C-0035

INTERLAMINAR SHEAR TESTING  
OF  
CARBON/CARBON COMPOSITES

by

William Kyros, Ph.D.

ABSTRACT

The goal of this effort is to define a shear test to determine accurately the interlaminar shear strength and shear modulus for 2-D, fabric reinforced, carbon/carbon advanced composites of thicknesses comparable to those envisioned for typical turbine engine components (about 1/4 inches thick and less).

A review of the literature on shear testing was performed and methods suitable for determining the interlaminar shear strength of thin, 2-D carbon/carbon materials were assessed. Methods for measuring shear strength and modulus on thick sections were assessed for their potential to be modified to the testing of thin sections on the order of 1/4 inch and less. The assessment considered torsion,  $10^{\circ}$  off-axis tensile,  $\pm 45^{\circ}$  balanced laminate tensile, rail shear, Iosipescu, slotted or double notched specimens, short beam bending and all tests analogous to those used in the testing of adhesives.

Two shear tests were developed which offer the best potential for satisfying the carbon/carbon shear test requirements. They are a torsion shear and an Iosipescu test. Fixtures for the Iosipescu test were designed as part of this effort.

### Acknowledgments

The technical support of Mr. Glenn G. Ormbrek (who served as the Effort Focal Point) and Dr. Nicholas J. Pagano both of the Air Force Materials Laboratory, Wright-Patterson Air Force Base, Ohio is gratefully acknowledged. Mr. George P. Peterson, Director of the Laboratory, and the Air Force Systems Command, Air Force Office of Scientific Research, are also acknowledged for providing the opportunity to conduct this study.

## OUTLINE

- I. INTRODUCTION
- II. GOALS AND OBJECTIVES
- III. APPROACH
- IV. RESULTS AND DISCUSSION
  - Shear Test Characteristics
  - Initial Screening
  - Torsion Shear Testing
  - Iosipescu Shear Testing
  - Lap Shear
  - Double Lap Shear
  - 10° Off-Axis Tensile Test
  - +45° Balanced Angle Laminate Under Tension (ASTM D3518-76)
  - Rail Shear Test
  - Double Notched
  - Plate Shear
  - Short Beam Shear Test (ASTM D2344-76)
  - Panel Shear Test
  - Asymmetric Four-Point Bending
- V. RECOMMENDATIONS AND PROPOSALS
  - Iosipescu Interlaminar Shear Test
  - Interlaminar Torsion Test

## REFERENCES



### LIST OF FIGURES

- Figure 1a. Inelastic Behavior in a Butt Joint  
Figure 1b. Torsion Tube Test  
Figure 2. Iosipescu Shear Test  
Figure 3. Bending, Shearing and Elongations in an Offset Flat Lap Joint  
Figure 4. Shear-Stress Concentrations in Double Lap Joints  
Figure 5.  $10^{\circ}$  Off-Axis Tensile Test  
Figure 6. Shear Forces in a Composite Specimen  
Figure 7. Applying Shear (In-plane or Interlaminar)  
Figure 8. Plate Shear Test  
Figure 9. Panel Shear Test  
Figure 10. Interlaminar Shear Test  
Figure 11. Interlaminar Torsion Test

### LIST OF TABLES

- TABLE I. MEASURABLE PROPERTIES AND STRESS STATES OF AVAILABLE SHEAR TEST SPECIMENS

## I. INTRODUCTION

This research essentially responds to a segment of a preliminary work statement developed by Dr. Nicholas J. Pagano, Materials Research Engineer, Air Force Materials Laboratory, Wright-Patterson AFB (Pagano, 1984). The work statement is from a request for proposals for the investigation of the shear properties of advanced carbon/carbon. It has as its objective the development and validation of a reliable test method to define interlaminar shear strength of advanced, 2-D, fabric reinforced, carbon/carbon composite materials. However, its scope is broad. It requires the review and assessment of existing material property data for a well-characterized 2-D, carbon/carbon material and a laminated structure constructed of this material which has already been subjected to stress analysis and tested to failure. It involves the design and testing of a control test specimen which reproduces the loading conditions of the failed laminated structure in an effort to precisely evaluate the interlaminar shear conditions that produced failure. A further requirement is that the contemporary approaches to determine material interlaminar shear strength be assessed from the viewpoints of specimen design optimization and further experimental inquiry leading to the development of a new interlaminar shear test. It is this latter requirement which this summer research effort addresses.

The rationale and justification for the development of a valid interlaminar shear test for carbon/carbon composites is largely based on the Pagano work statement and on preliminary design studies that are cited below.

2-D, fabric reinforced, carbon/carbon composite materials show considerable promise for turbine engine component applications such as augmentor flaps and seals, blades, hubs, shrouds for axial and radial flow turbine engine rotors, combustors, and combustor liners. However, a credible database of allowable properties for structural design does not yet exist for the proposed materials (Cruzen, 1980; Forcht et al, 1981; Schmid, 1982). The ELITE program currently underway is intended to improve the database for design purposes as well as to support the materials development activity. In addition, preliminary design studies show that thin sections of 2-D, fabric reinforced, carbon/carbon materials show negative margins under shear loading conditions. This coupled with an inability to accurately define interlaminar shear allowables severely limits the ability to design reliable 2-D laminated carbon/carbon, thin-walled structures for turbine engine applications.

Therefore, a valid test method is needed to determine the interlaminar shear strength of thin, high modulus, 2-D, fabric reinforced, carbon/carbon composites. Such a test will provide a quantitative basis to address the following critical design related issues more effectively: (1) the more reliable prediction of composite failure in service, (2) the more rational control of material processing variables, (3) the more rational selection of constituent materials with respect to their geometries and properties, and (4) the effects of oxidation protection on the mechanical performance of the carbon/carbon.

## II. GOAL & OBJECTIVES

The goal of this effort is to define a shear test to determine accurately the interlaminar shear strength and shear modulus for design purposes for 2-D, fabric reinforced, carbon/carbon advanced composites of thicknesses comparable to those envisioned for typical turbine engine components (about 1/4 inches thick and less).

More specifically, the objectives were:

1. To delineate and rationalize all of the desirable characteristics for an acceptable shear test for the general class of 2-D carbon/carbon composites to generate engineering design data.
2. To generally review all extant shear tests and the extent to which they possess the desirable characteristics.
3. To select the "best" tests and modify them to maximize their effectiveness in generating accurate, reproducible shear data with acceptable measures of dispersion.

## III. APPROACH

A review of the literature on shear testing was performed and methods suitable for determining the interlaminar shear strength of thin, 2-D carbon/carbon materials assessed. Methods for measuring shear strength and modulus on thick sections were assessed for their potential to be modified to the testing of thin sections on the order of 1/4 inch and less. The assessment considered torsion,  $10^{\circ}$  off-axis tensile,  $\pm 45^{\circ}$  balanced laminate tensile, rail shear, Iosipescu, slotted or double notched specimens, short beam bending, and all tests analogous to those used in the testing of adhesives. A list of references is located at the end of this report.

Existing shear data on all 2-D, fabric reinforced, carbon/carbon composites was reviewed and analyzed for reproducibility, accuracy, and to determine differences in shear values obtained from different tests on the same material. Stress analyses performed on test specimens was examined and used to develop more promising specimen configurations. Specimen configurations were analyzed and modified to produce sufficient strains in gage sections to ensure the performance of valid tests and accurate strain measurements.

Based upon the literature reviews, existing interlaminar shear data, and specimen stress analyses, new shear test methods were developed and proposed. It is noted that while the determination of accurate interlaminar shear strength is the primary objective of the new shear test, the simultaneous measurement of shear modulus was considered essential.

#### IV. RESULTS AND DISCUSSION

The importance of accurately measuring interlaminar shear properties of carbon/carbon materials for design purposes has been established. And although the literature reveals several methods to determine shear properties on composite materials, not all have been developed or are useful for the generation of design information and therefore are not valid for this purpose. Since the goal of this work is to develop a method of evaluating interlaminar shear allowables and moduli for design purposes for the general class of 2-D fabric-reinforced carbon/carbon composites, the ideal shear test is required to exhibit those most desirable characteristics that will generate such data on these particular materials with their idiosyncrasies.

The characteristics upon which the assessment of the variety of shear tests previously cited are delineated and rationalized below to represent the most ideal features of a test that can meet the data requirements dictated by the turbine engine applications. This section discusses these characteristics, the assessment of the extant tests, and the selection of the most promising methods.

##### Shear Test Characteristics

It should be noted that the characteristics discussed here are best described as "most desirable" and not as "essential" features, for it is clear at the outset that no single test will possess all of the attributes necessary to generate the kind of shear design data required. In order to develop such

a test, the specimen configuration and test fixtures will have to be "engineered" to an acceptable product. That is, desirable features, essential attributes and data purity must be synthesized, analyzed, and alternatives considered to produce an acceptable compromise design. Obviously certain characteristics ought to be weighed more than others. For example, in order to generate shear allowable data, greater importance should be placed on developing a pure shear stress state in the test section rather than the desirable characteristics of ease of testing and small specimens of simple geometry. In short, a scientific breakthrough in specimen design is not likely to yield a solution to the interlaminar shear testing problem. Instead, the essential features and desirable characteristics ought to be optimized using an analytical-experimental optimization process.

The desirable and essential characteristics of the test are as follows:

1. Uniform stress distribution in the test section
2. Pure shear stress state in the test section
3. Minimization of edge effects (by consideration of coupling phenomena)
4. Strains measurement on macro-scale via clip on extensometer
5. Interlaminar strength and modulus.
6. Specimen size
  - a. small enough to be economical
  - b. thin (less than 0.2 inches or less)
  - c. simple geometry
  - d. large enough so that defects reflect an "organized" response
7. Controlled testing rates
8. Simple and inexpensive fixtures
9. Ease of testing

#### Initial Screening

Table I summarizes the measurable properties and stress states of available shear test specimens.

#### Torsion Shear Testing

Torsion of a thin wall tube creates a state of near-uniform pure shear in the tube wall, and the angle of twist per unit length gives a measure of the shear strain. The main disadvantages of this method are the limitations imposed by the specimen geometry. Only specimens in the form of a tube whose axis is parallel to the axis of symmetry of the material give uniform

TABLE I.  
MEASURABLE PROPERTIES AND STRESS STATES  
OF AVAILABLE SHEAR TEST SPECIMENS

METHOD	PROPERTIES	STRESS STATE	EDGE EFFECTS	REFERENCES
Torsion Solid X Section	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Pure shear Variable distrib.	Minor	Causey, 1981 Perry, 1959 Hancox, 1970
Torsion Thin Walled X Section	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Pure shear Uniform distrib.	Minor	N/A
Iosipescu	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Pure shear Variable distrib.	Minor	Herakovich, 1981 Iosipescu, 1967
Lap Shear	$\tau_{ULT}$ $\sigma_{ULT}$	Shear and normal Variable distrib.	Severe	Perry, 1959
Double Lap Shear	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear and normal Variable distrib.	Severe	Perry, 1959 Ishai, 1977
10° Off-Axis Tensile	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear and normal Variable distrib.	Minor to Severe	Chamis, 1977 Sinclair, 1977
± 45° Balanced Laminate Tensile	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear and normal Uniform distrib.	Minor	Whitney, 1982 Hahn, 1973
Rail Shear	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Pure shear at ctr. Uniform distrib.	Severe	Terry, 1979 Whitney, 1971
Double Notched	$\tau_{ULT}$	Shear and normal Variable distrib.	Severe	Hancox, 1970 Herakovich, 1981
Plate Shear	$G$	Shear Uniform distrib.	None	Whitney, 1967
Short Beam	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear and normal Variable distrib.	Severe	Hancox, 1970
Panel Shear	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear and normal Variable distrib.	Severe	N/A
Asymmetric Four-Point Bending	$\tau_{ULT}$ $\sigma_{ULT}$ $G$	Shear Variable distrib.	Minor	Iosipescu, 1967 Slepetz, 1978 Walrath & Adams, 1983

distribution of shear stress. These geometrical limitations make this specimen geometry difficult for use on the application. Torsion of solid cross-section specimens is a much more feasible and leads to results consistent with other approaches (Causey 1981, Perry 1959, Hancox 1970). This test is adaptable as an interlaminar shear test. See Figures 1a. and 1b.

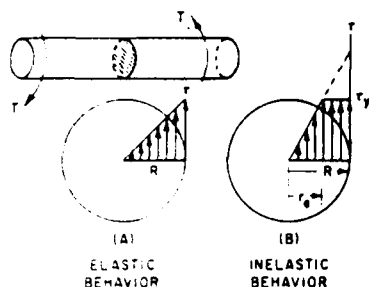


Figure 1a. Inelastic Behavior in a Butt Joint

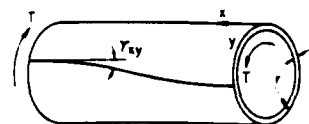


Figure 1b. Torsion Tube Test

### Iosipescu Shear Testing

The Iosipescu shear test is illustrated in Figure 2. This test can be used to measure the in-plane shear strength of 1-D or 2-D composites. With rigid fixtures the stress state in the center of the gauge section is close to uniform shear, but at the notch tip a significant normal tensile stress may be present, (Herakovich 1981, Iosipescu 1967). The literature is unclear about the extent of the normal stress distribution and this should be investigated further. This test is adaptable as an interlaminar shear test.

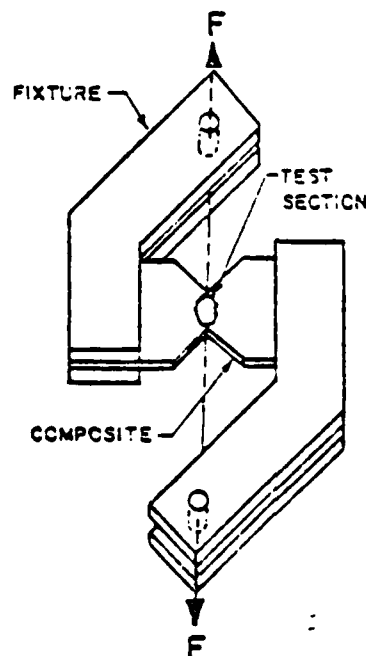


Figure 2. Iosipescu Shear Test  
83-11

### Lap Shear

Flat lap joints in tension (Figure 3.) have been the subject of considerable analysis (Perry 1959). Of these the theory of Goland and Reissner (Goland and Reissner 1944) seems to be the most useful in that it allows for bending as well as differential shearing.

The Goland-Reissner theory permits the calculation of the shear stress  $\tau$  and the tensile stress  $\sigma$  throughout the joint. However, the strength of the joint is limited by the greatest stress on the adhesive which occurs at the ends of the laps.

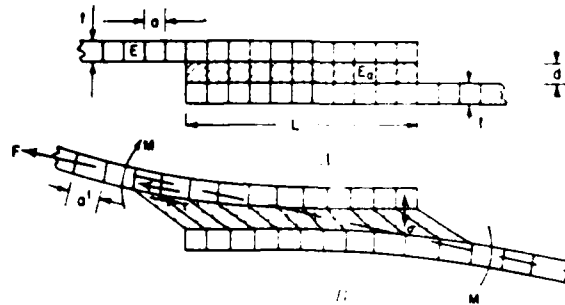


Figure 3. Bending, Shearing and Elongations in an Offset Flat Lap Joint

### Double Lap Shear

The stress-concentration factors for double lap joints have been derived by Volkersen (1938). The formula and a useful set of curves are given in Figure 4. This theory appears to be applicable to flat lap joints under tension and compression, if bending is prevented and hence is valid for double laps in tension or compression if they are symmetrical on either side of the plane action. This requirement for validity of the theory in double lap joints is that the adherends and glue lines must be of like materials  $E_2$  of equal thicknesses  $t_2$ . However,  $E_1$  may differ from  $E_2$  and  $t_1$  may differ from  $t_2$ . Typical calculated values of the maximum shear stress at the free boundary versus parameter  $\Delta$  as a function of the relative stiffness parameter  $W$  are given in Figure 4.

Perry (1959) and Ishai (1977) have reported the potential of this test. With proper instrumentation for the measurement of interlaminar strain and with considerable care to minimize the effects of bending, the double lap shear offers some potential and should be investigated in the context of this application.



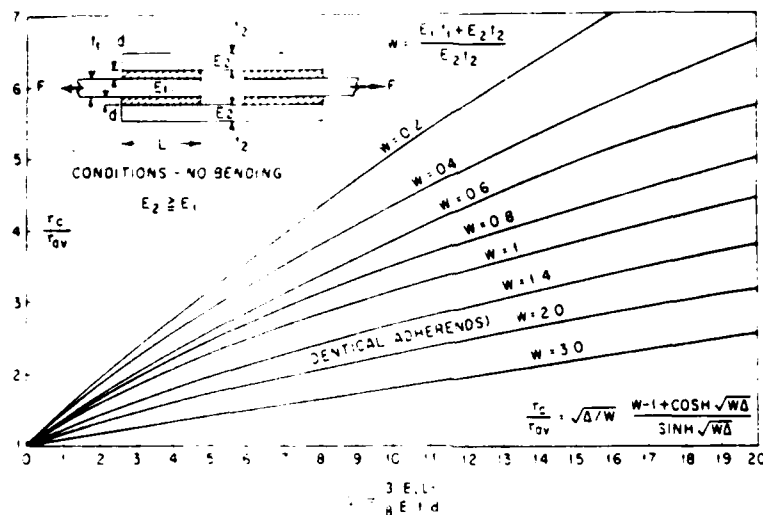


Figure 4. Shear-Stress Concentrations in Double Lap Joints

#### 10<sup>0</sup> Off-Axis Tensile Test

An off-axis tensile test can be used to measure the in-plane shear strength of a unidirectional composite (Chamis, Sinclair 1977). The stress state is shown in Figure 5. In addition to shear stress, an in-plane tensile stress parallel to the fibers of 5.7 times the shear stress and an in-plane tensile stress transverse to the fibers of 0.2 times the shear stress exist. Thus, the stress state is not pure shear and failure may be influenced by these tensile stresses. In a weakly-bonded composite such as carbon/carbon, failure may even occur through delamination under the transverse stress. The shear strain must be measured by applying a strain gauge rosette to the specimen and transforming the strains.

The main advantages of this test method are the simple specimen geometry and the tensile loading mode, although clamped end must be avoided because they can cause large bending moments. The main disadvantage, in addition to the complex stress state described above, is the long gauge length-to-width ratio (14:1) required and the correspondingly large dimensions of the parent material.

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

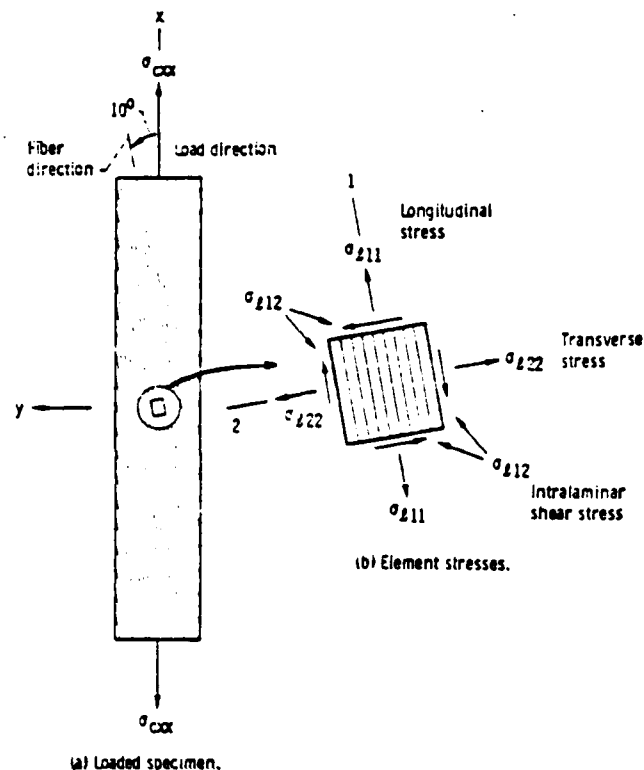


Figure 5. 10° Off-Axis Tensile Test

#### ± 45° Balanced Angle Laminate Under Tension (ASTM D3518-76)

If a [45/-45/45/-45]<sub>s</sub> laminate is tested in tension or compression, the laminae are subjected to an in-plane shear stress equal to half the nominal in-plane tensile or compressive stress, combined with a normal tensile or compressive component (Whitney 1982). Even though the shear characteristics may be influenced by the presence of the normal stress, the magnitude of this influence can be estimated by performing tests in both tension and compression on the same laminate. Strain gauge rosettes are required to measure the shear strains and the laminate needs to be laid up accurately to ensure a valid test (Hahn 1973).

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

### Rail Shear Test

In this test, shear forces are applied to a rectangular composite specimen through a rigid steel rails bolted or bonded to the edges, Figure 6. If the rails are ideally rigid, the stress state is close to uniform pure shear, but non-rigid rails cause some stress concentration and introduce a normal stress component (Terry 1979, Whitney 1971). A strain gauge rosette is required for accurate strain determinations; approximate strains can be estimated from the loading point displacements.

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

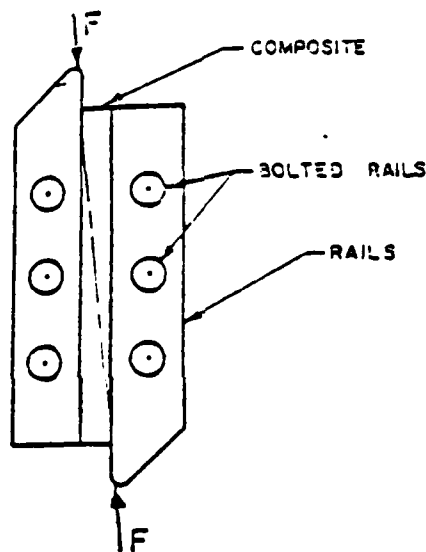


Figure 6. Shear Forces in a Composite Specimen

### Double Notched

This test method offers a simple means of applying in-plane shear to 1-D or 2-D composites, or interlaminar shear to 2-D composites, Figure 7. The specimens can be small and the fixtures simple. The disadvantages are the presence of large stress concentrations and high normal stresses, (Herakovich 1981, Hancox 1970).

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

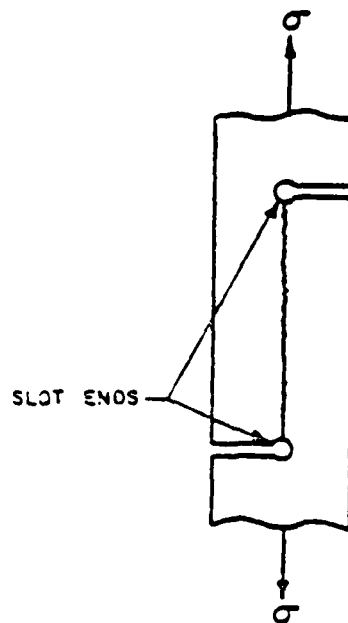


Figure 7. Applying Shear (In-plane or Interlaminar)

#### Plate Shear

The plate test is shown in Figure 8. If the deflection is measured along the diagonals, and the origin of the coordinates is assumed to be fixed, then, at a diagonal distance,  $d$ , from the origin, the transverse deflection is

$$W = 3Pd^2/2h^3G$$

Whitney (1967) discussed the limitations of this test on the basis of the assumption of homogeneity and orthotropy.

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

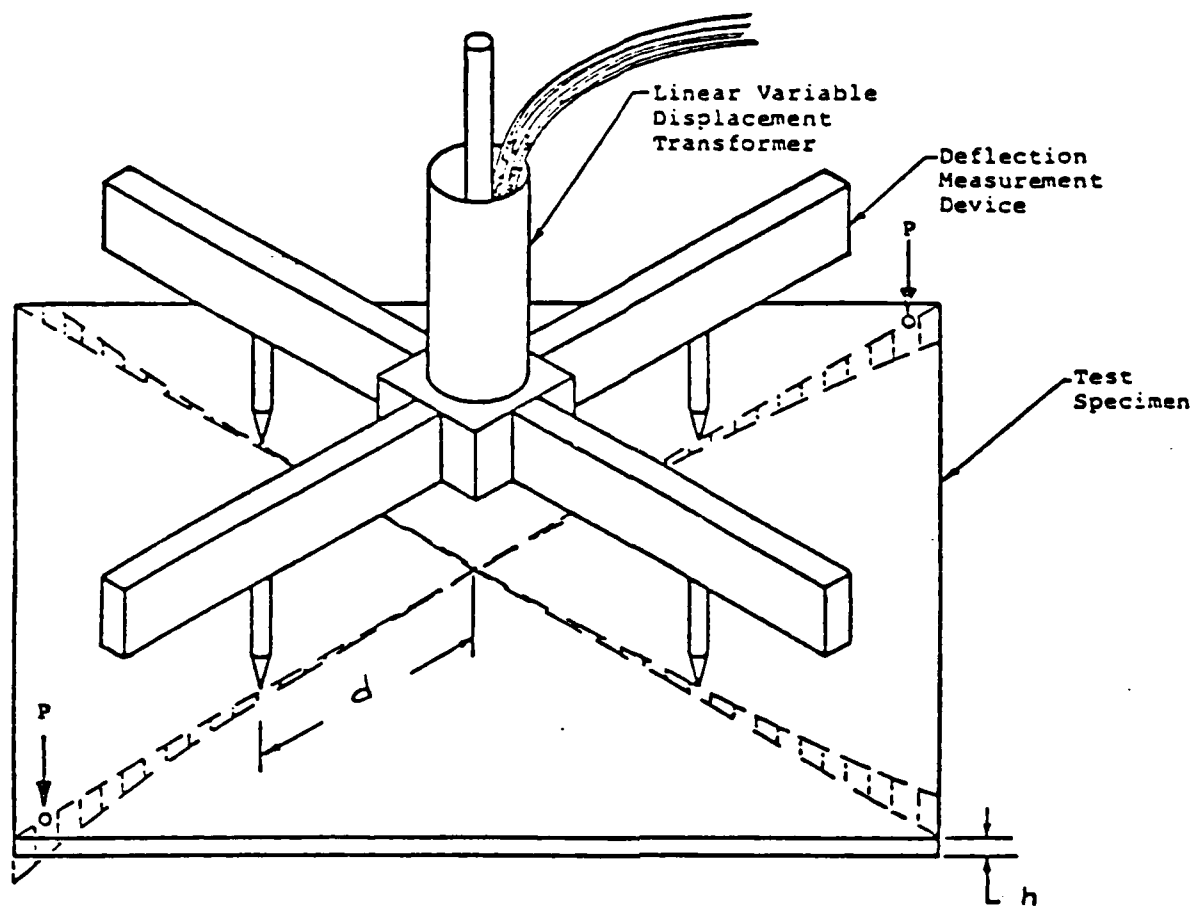


Figure 8. Plate Shear Test

#### Short Beam Shear Test (ASTM D2344-76)

This method can be used to measure the interlaminar shear strength of 1-D and 2-D composites only. A short beam with a low span-to-depth ratio is tested in 3-point loading. By making the span-to-depth ratio low enough, the beam can be made to fail in shear rather than flexure. The shear stress varies with position in the beam, and its value must be calculated from homogeneous beam theory.

Because of the uncertainties associated with calculating the shear stress and the influence of normal stresses, the short beam test does not seem to offer much advantage over methods such as the double shear method, (Hancox 1970).

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

### Panel Shear Test

Figure 9. shows the test fixture and specimen. Panel shear specimens are tested in the four-link rig shown schematically. Shear is introduced to the specimen from calibrated, adjustable loading links via steel edge members bolted to each side of the specimen reinforcement. It is necessary to inhibit specimen buckling using Teflon-coated wooden blocks clamped to each side of the test area, independent of the edge members and clear of the specimen reinforcement.

This test is not considered a good candidate for an interlaminar shear test for this application to carbon/carbon materials.

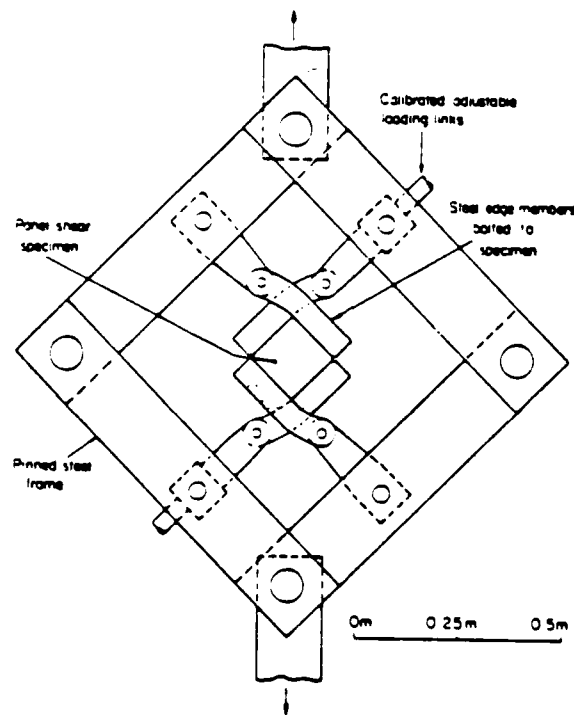


Figure 9. Panel Shear Test

### Asymmetric Four-Point Bending

Based upon Iosipescu's original work, Iosipescu (1967), several modified versions of the test method have evolved. Slepetz, utilized a slightly modified loading scheme, and termed the test the "asymmetrical four-point bending (AFPB) test." While this modification permits easier specimen

loading, the induced shear stress becomes a function of the loading-point location dimensions, a distinct disadvantage in comparison to the Iosipescu configuration. Slepetz did do a very thorough study, however, including an investigation of stress uniformity using strain gages, moire-fringe interference, and a finite-element analysis. The latter was also used to study the influence of specimen notch geometry (Slepetz 1978; Walrath and Adams 1983).

This type of test does offer some potential and should be considered in the same light as the Iosipescu Test.

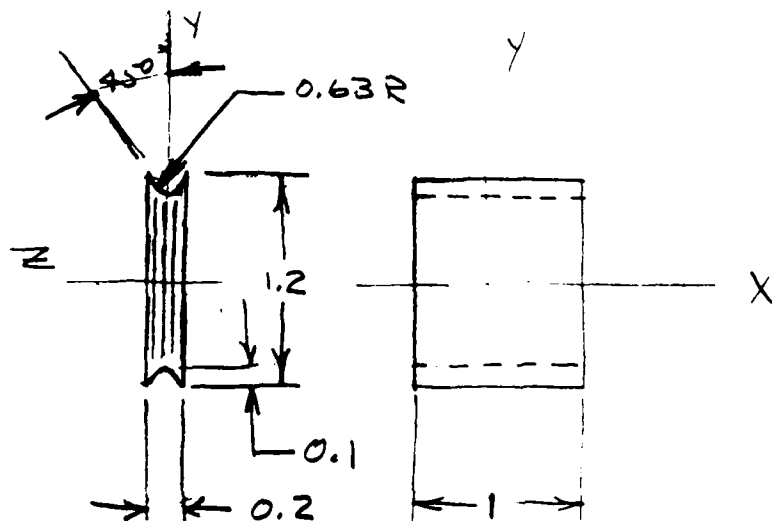
V. RECOMMENDATIONS AND PROPOSALS

IOSIPESCU INTERLAMINAR SHEAR TEST

Utilize Iosipescu test fixture with specimen oriented so that applied shear stress will coincide with interlaminar planes.

Requires specimen be bonded to steel studs.

COMPOSITE SPECIMEN



TEST SPECIMEN

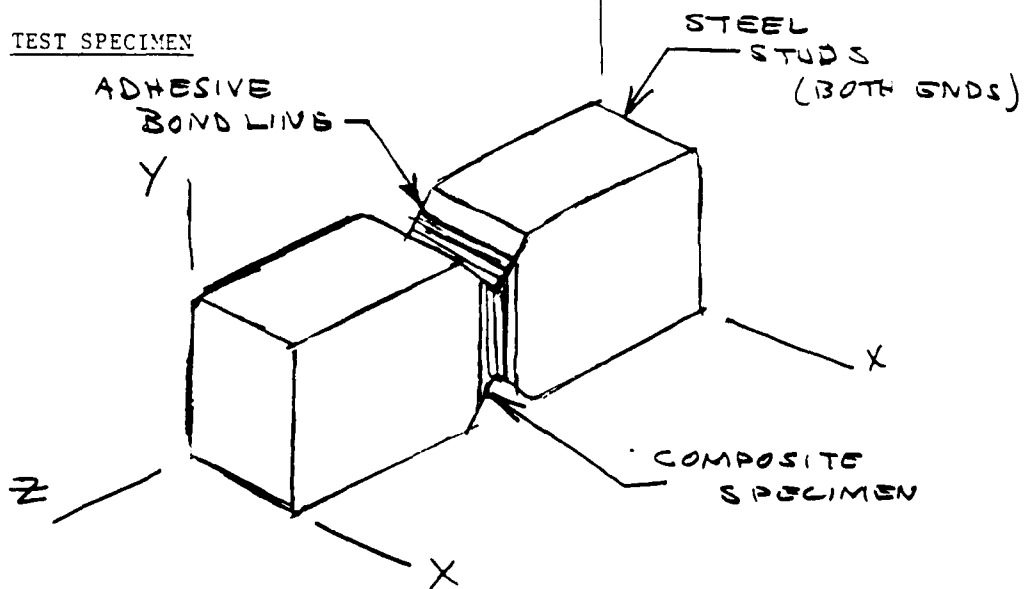


Figure 10. Interlaminar Shear Test  
83-20



#### ADVANTAGES

Almost pure shear stress state edge effects and coupling phenomena minimized as a result of no inplane extensional or bending strains.

Both interlaminar strength and modulus can be measured.

#### PROPOSED STUDY

Design and construct Iosipescu and Asymmetrical Four-Point Loading fixtures and test various materials; specimen configurations to determine effects of notch geometry.

Perform finite element analyses on specimen configurations and compare with experimental data.

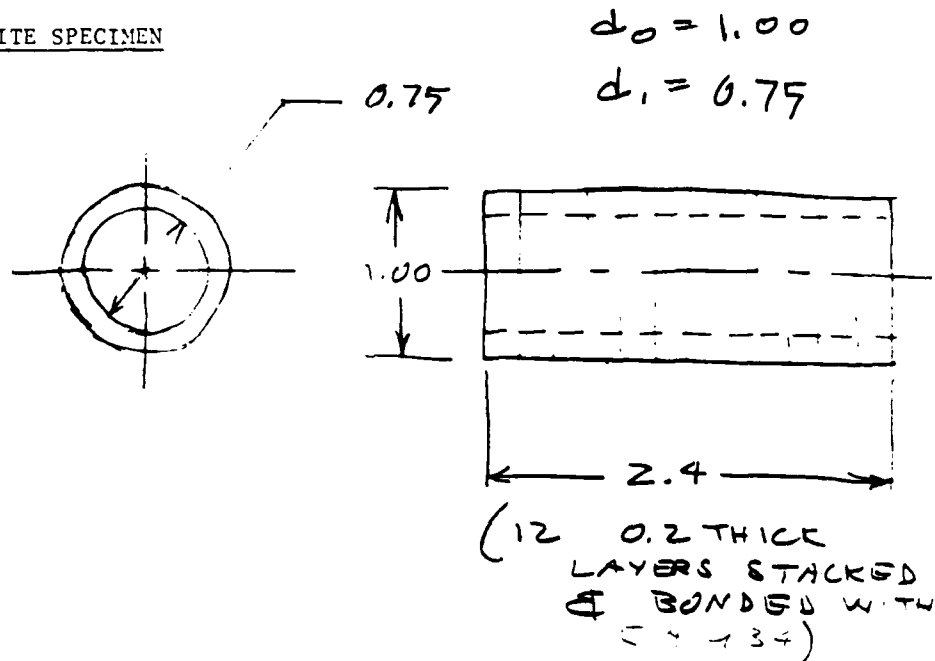
Evaluate test fixtures, tension versus compression, tolerance requirements, ease of testing, variation of data, and failure modes.

### INTERLAMINAR TORSION TEST

Apply torsion to ring specimen configuration of carbon/carbon composite so that applied shear stress coincides with the interlaminar planes.

Require material to be stacked to thickness which will produce sufficient shear strain to enable the accurate measurement of shear modulus as well as strength.

#### COMPOSITE SPECIMEN



#### TEST SPECIMEN

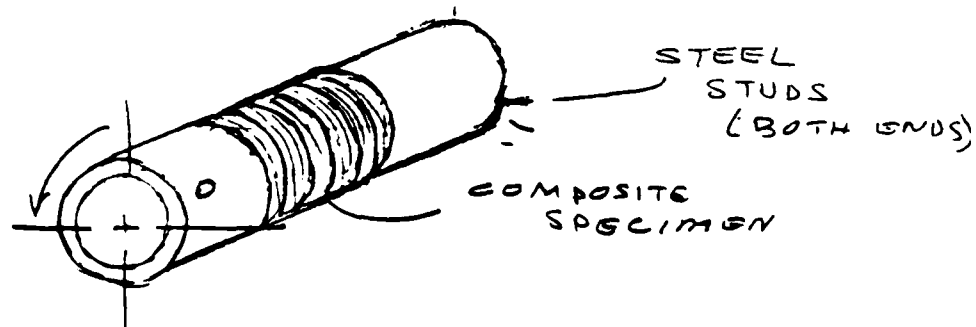


Figure 11. Interlaminar Torsion Test

## REFERENCES

- Causey, S.J., (1981). "A Discussion on the Properties of Several Candidate MX Rocket Nozzle Materials Obtained from Ring Tests." CPIA Publication 346. January.
- Chamis, C.C., and J.H. Sinclair, (1977). "10° Off-Axis Test for Shear for Properties in Fiber Composites." NASA Technical Memorandum NASA TM X-73550. April.
- Cruzen, G.S., et al, (1980). "Advanced Gas Turbine Carbon-Carbon Components Feasibility Assessment." AFWAL-TR-80-4036. April.
- Forcht, B.A., et al, (1981). "Carbon-Carbon Technology for Future Strategic and Tactical Missiles." Three Volumes. AFWAL-TR-82-4131. November.
- Goland, M. and E. Reissner, (1944). "The Stresses in Cemented Joints." Journal of Applied Mechanics. March.
- Hahn, H.T., (1973). "A Note on the Determination of the Shear Stress Strain Response of Unidirectional Composites." Journal of Composite Materials. July.
- Hancox, N.L. and H. Wells, (1970). "A Comparison of Three Methods of Measuring Interlaminar Shear Strength of Unidirectional CFRP." Carbon Composite Technology Symposium, ASME. January.
- Herakovich, C.T., et al., (1981) "A Comparative Study of Composite Shear Specimens Using the Finite Element Method." Test Methods and Design Allowables for Fibrous Composites. ASTM STP-734, C.C. Chamis EP. pp.129-151.
- Iosipescu, N., (1967). "New Accurate Procedure for Single Shear Testing of Metals." Journal of Materials. Volume 2, Number 3. September.

- Ishai, O. et al., (1977). "Direct Determination of Interlaminar Stresses in Polymeric Adhesive Layer." Experimental Mechanics. July.
- Pagano, N.J., (1984). "Shear Properties of Advanced Carbon-Carbon." Unpublished communication. Description/Specification for Purchase Request.
- Perry, H.A., (1959). Adhesive Bonding of Reinforced Plastics. New York: McGraw-Hill.
- Schmid, T.E., (1982). "Oxidation Resistant Carbon-Carbon Composites for Turbine Engine Aft Sections." AFWAL-TR-82-4159. October.
- Slepetz et al., (1978). "In-Plane Shear Test for Composite Materials." Army Materials and Mechanics Research Center, Watertown, Massachusetts. Technical Report AMMRC TR78-30. July.
- Terry, G., (1979). "A Comparative Investigation of Some Methods of Unidirectional In-Plane Shear Characterization of Composite Materials." Composites, Page 233. October.
- Volkersen, O., (1938). Luftfahrt-Forsch 15:41.
- Walrath, D.E., and D.F. Adams, (1983). "The Iosipescu Shear Test as Applied to Composite Materials. Experimental Mechanics. March.
- Whitney, J.M., (1967). "Experimental Determination of Shear Modulus of Laminated Fiber Reinforced Composites." Experimental Mechanics. October.
- Whitney, J.M., (1971). "Analyses of Rail Shear Test Applications and Limitations." Journal of Composite Materials. January.
- Whitney, J.M., et al, (1982). "Experimental Mechanics of Fiber Reinforced Composite Materials." Society for Experimental Stress Analysis Monograph No. 4.

1984 USAF-SCEE SUMMER FACULTY RESEARCH PROGRAM  
Sponsored by the  
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH  
Conducted by the  
SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION  
FINAL REPORT  
TWO-DIMENSIONAL MEDIAN FILTERS FOR IMAGE  
PREPROCESSING IN MACHINE RECOGNITION

Prepared by:	Dr. David C. Lai
Academic Rank:	Professor
Department and University:	Department of Electrical Engineering and Computer Science, The University of Vermont
Research Location:	Rome Air Development Center Intelligence & Reconnaissance Division Image Systems Branch, Image Exploitation Section
USAF Research:	John Potenza
Date:	20 July 1984
Contract No:	F49620-82-C-0035

TWO-DIMENSIONAL MEDIAN FILTERS FOR IMAGE  
PREPROCESSING IN MACHINE RECOGNITION

by

David C. Lai

ABSTRACT

The aim of this project is to explore and evaluate the effectiveness of median filtering and median/inverse filtering in facilitating machine recognition, and to gain insight into the design of practical preprocessors for automatic pattern recognition systems. 2-D median filters of various shapes and sizes are first evaluated and compared based on their performances in machine recognition under a particular degradation condition. Based on this result, the median filter with the best shape and size was selected and used throughout the experiment for measuring its performances in machine classification under various degradation conditions. This 2-D median filter was also used in conjunction with inverse filters to form preprocessors. Their performances in enhancing automatic pattern recognition were measured and compared. Results are reported and suggestions for future research are given.

#### ACKNOWLEDGEMENT

The author wishes to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him the opportunity to spend an interesting summer at Rome Air Development Center, Griffiss AFB, N.Y. He would also like to acknowledge the hospitality and good working conditions provided by the Image Exploitation Section at RADC.

This work would not have been possible without the support of many individuals at RADC. Many gave up their computer time to accommodate our project. In particular, the author would like to thank Mr. John Potenza, his USAF research colleague, for his collaboration, suggestions and fruitful discussions, and Mr. Frederick W. Rahrig, who was always willing to lend a helping hand on the computing facility, for his suggestions in selecting test images, features and classifier.

Last but not least, he wishes to thank Mr. Kevin Verfaille of the University of Vermont for many contributions to this project and faithfully carrying out the experimental work.

## I. INTRODUCTION:

Median filters are non-linear filters. Median filtering was first introduced by Tukey<sup>1</sup> for smoothing time series, wherein he noticed its property in preserving large sudden changes of level (which are equivalent to the edges in images) in time series. It has later been adapted for use in speech processing.<sup>2,3,4</sup> In that, they noted that the median filtering preserves discontinuities of sufficient duration and yet eliminates local roughness in the speech signal and performed a study to examine the effect of median filtering in the suppression of impulse noise introduced by bit errors in the transmission of digital speech signals.

Median filtering has also been adapted for use in image processing.<sup>5,6,7,8</sup> Pratt<sup>5</sup> made a very qualitative study of 2-D median filters of two sizes and shapes and applied 1-D median filters to a picture contaminated by impulse noise. Huang<sup>7</sup> reviewed the use of median filters in image processing and showed, by examples, the effect of window size and shape on processing images corrupted by impulse noise and others used median filtering to remove scanner noise.

Because of the difficulty involved in the theoretical analysis of 2-D median filters, published results on the characteristics and theoretical behavior of 2-D median filters are rare.<sup>9,10,11</sup> There are more extensive results on the properties of 1-D median filters reported.<sup>12,13,14,15,16</sup> However, the 1-D results cannot be readily extended to 2-D median filters. Justusson<sup>9</sup> studied various properties of the outputs of 2-D median filters with various signal and noise inputs such as Gaussian noise and impulse-type noise (also known as salt-and-pepper noise) and a step signal plus noise. Tyan<sup>10</sup> made a study on the fixed-point set of median filter (a fixed-point set is an array that is unmodified by median filtering). Nodes and Gallagher<sup>11</sup> derived the structures of fixed-point sets of the 2-D separable median filters. Ataman<sup>12</sup> compared the performance of 1-D median filtering on removing impulse-type noise and Gaussian white noise with that of Hanning average filtering. The others dealt, more or less, with fixed-point sets of 1-D median filters.



Although the theoretical behavior of 2-D median filters is not well understood, its simplicity in concept and design has lead to the development of fast algorithms for off-line or on-line image processing,<sup>17,18,19</sup> and schemes for hardware or VLSI implementation for real-time image processing.<sup>20,21,22,23,24,25</sup> Processing speed is a major concern in the design and use of a preprocessor in an automatic pattern recognition system; speed is not a crucial factor for preprocessing for human observers. Hence, the potential for VLSI or VHSI implementation of median filters for real-time image preprocessing is a great asset for automatic target identification and scene classification by machine recognition; particularly for use as preprocessor to Air Force automatic target identification and classification systems using reconnaissance imagery. As demonstrated in Refs. 5, 6, 7 and 8, there is ample evidence supporting the usefulness of median filters in suppressing impulse-type noise, Gaussian noise, scanner noise, and oscillating structures without smearing the edges as judged by human observers. However, the effectiveness of median filters as preprocessors for machine recognition systems has not been fully explored and evaluated. The effects of median filtering on machine analysis of images for edge detection, shape analysis, and texture analysis have been reported;<sup>26</sup> in that, they compared the effectiveness of median filters, variant of median filters, and moving average filters in extracting the edge and shape feature. Their results are encouraging; however, the effectiveness in extracting features does not necessarily guarantee effectiveness in classification. The effectiveness of image restoration filters in machine recognition has been evaluated;<sup>27</sup> in which, they evaluated and compared the performances of five linear filters, and a median filter in an automatic pattern recognition system for images degraded by Gaussian blur and additive Gaussian white noise with various degrees of severity. The performance of median filtering is promising but not conclusive. In order to use median filter as preprocessor for automatic pattern recognition system, we must not only explore and evaluate its effectiveness alone but also the usefulness of combining it with other types of filters.

Unlike the design of the linear filters, the median filters were not designed to restore images distorted by blurring. Median filters were designed mainly for removing noise while preserving sharp edges without any regard to restore a blurred image to its original form. The inverse filters, on the other hand, were designed to restore images degraded by blurs which can be characterized by PSF's (point-spread functions); however, inverse filters do not work well under noisy conditions.<sup>28</sup> It seems that the characteristics of inverse filters and median filters complement each other. It is natural then to use a median filter in tandem with an inverse filter so that a blurred image corrupted by noise may be restored. Since both of these filters are relatively simple to implement, the combination will serve as a practical preprocessor for any pattern recognition system. Again, the utility and effectiveness of this combination of inverse filtering and median filtering have not been assessed and evaluated.

This project concentrates on the investigation and evaluation of performance of median filters of five shapes in various sizes and the combination of a median and an inverse filter operated in tandem. Since automatic classification and identification are emphasized by the Air Force, we chose to evaluate filters based on how well they perform in facilitating machine classification rather than human interpretation. Since both median and inverse filters are simple to implement, it is hoped that the result will lead to the development of powerful practical preprocessors for use in automatic pattern recognition systems.

The objectives of this project are reiterated in the next section. A section on the properties of median filters and inverse filters will then follow. The experiment for evaluation and comparison of median filters and median/inverse filters operating under various noise and blur conditions is described in a section. Experimental results and their interpretation are given in another section. Finally, recommendations for further research will be listed.

## II. OBJECTIVES:

The goal of this project is to explore, analyze, evaluate and compare the performances of 2-D median filters of various shapes and sizes and the combination of median and inverse filters in machine recognition. To be more specific, we list the following objectives:

(1) To implement one of the fast median filtering algorithms as a software package to expand RADC's "IRAMS" capability.

(2) To analyze median filtering mainly to gain insights for their application in facilitating machine recognition.

(3) To test median filters of various shapes and sizes alone and then in tandem with inverse filters on images degraded by various types of noises and blurs in various degrees of severity so as to evaluate and compare their effectiveness as preprocessor for an automatic pattern recognition system.

(4) To develop guidelines for selecting the filter window shape and size so that it will operate most effectively under a given degradation condition.

This project has not been concerned with the interrelationship between filters (preprocessors), feature extractors, and classifiers. It was also unable to encompass a larger class of window shapes though the filter shapes selected for evaluation in this project are representative. It was also not feasible to test each median filter (of a definite shape and size) against a large number of possible degradation conditions.

## III. PROPERTIES OF 2-D MEDIAN FILTERS AND INVERSE FILTERS:

The output of a non-recursive median filter is given as

$$m(x,y) \triangleq \underset{w}{\text{median}}[g(x,y)] \triangleq \text{median}[g(x+r, y+s); (r,s) \in w],$$

$(x,y) \in \text{input image data set}$  (1)

where  $w$  is the filter window which is usually centered at the pixel located at  $(x, y)$ ;  $g(x, y)$  denotes the value of the input pixel located at  $(x, y)$ ; and  $g(r, s)$ , an input pixel inside the window. The coordinate variables  $r$  and  $s$  are varied to cover all the pixels inside the window and  $x$  and  $y$  are varied to move the windows to cover another portion of the image. The filter window may assume many shapes such as line segments (reduced to 1-D), rectangles, squares, circles, crosses, slanted crosses, square frames, etc. of various sizes. For border points, the median is usually computed on those points covered only by the window (i.e. the window should not be filled up by padding zeros). In general, the number of pixels inside the window is odd; otherwise, the mean of the two middle pixel values is taken as the median value. As the window moves on, the center pixel value  $g(x, y)$  of the window is replaced by the median pixel value  $m(x, y)$  just computed; this median pixel value will not be used for computation of the median of pixels in the next window. That is, the window covers only the original pixels of a given image, thus the name non-recursive median filtering. In practice, several median filtering passes over the image may be needed to obtain good results.

Modifications and generalization of median filtering as described above have also been used or proposed by other researchers.<sup>29, 30, 31</sup> Lehar and Stevens<sup>29</sup> obtained good results by using a thresholded median filtering in which the center pixel of the window is not replaced by the median value of the pixels inside the window unless the difference between it and the window median exceeds a preset threshold. Bovik et al<sup>30</sup> offered a scheme for designing the order-statistic filters which enables one to design a linearly combined order-statistic filter optimized in the least-mean-square sense. Bendnar and Watt<sup>31</sup> compared the  $\alpha$ -trimmed means and median filters in which the  $\alpha$ -trimmed mean filter becomes median filter when  $\alpha$  reaches 0.5; it thus offered some insight about median filters. These variations of median filtering are mentioned here for the sake of completeness but they are beyond the scope of this report.

The design of a median filter rests only on the shape and size; i.e., a 2-D median filter is uniquely determined by its shape and size. Once the filter window is selected, the filtering operation or computation is performed by following Eq. (1). However, there exists no procedures or rules for selecting appropriate window shape and size for use under a specified operating condition. It is hoped that our results may shed some light on how to select the window shape and size.

Statistical properties of median filters as derived from theoretical work<sup>9,12</sup> confirmed, in a more quantitative manner, the ability of median filters in removing noise while preserving edges. The theoretical comparison were, in general, performed between median filters and moving average filters. It was concluded that median filters are more effective in removing impulse-type noise or noise with heavy-tailed distributions than moving average filters while also preserving edges better. In removing Gaussian white noise, the moving average filters are superior, however they smear edges.

It was shown<sup>9</sup> that, for an input signal which constitutes an image signal constant in the window with additive zero-mean white noise of symmetric p.d.f., the median-filter output  $m(x, y) = \text{med}_w g(x, y)$  has a distribution which is approximately Gaussian for large  $K$ , the size of the window. The mean and variance approach asymptotically to

$$\begin{aligned} E[m] &\simeq \text{the constant signal level} \\ \text{Var}[m] &\simeq \frac{1}{4K\rho^2(0)} \end{aligned} \quad (2)$$

where  $p(n)$  is the noise p.d.f. If the white noise is Gaussian distributed, Eq. (2) becomes

$$\text{Var}[m] \simeq \frac{\pi}{2} \frac{\sigma_n^2}{K}$$

where  $\sigma_n^2$  represents the Gaussian white noise power. From above, we see that the amount of noise reduced is inversely proportional to the window size; i.e., the larger the window size, the more effective the noise reduction, however the resultant image will be more distorted. Therefore, in selecting window size for a median filter, one has to contend with the trade-off between noise reduction and image distortion. As a matter of fact, when the window size is too large, certain shapes or features in the image can be eliminated completely. In this study the performance of machine classification is used as the criterion for selecting window size and shape and provides one with a glimpse of this trading-off effect on machine classification.

The median filtering is a nonlinear operation though it is known that

$$\text{med}_w [\alpha g(x, y) + \beta] = \alpha \text{med}_w [g(x, y)] + \beta$$

where  $\alpha$  and  $\beta$  are two constants; i.e., it has the appearance of linearity under scaling and/or translation. If the input pixels to the median filter is scaled and/or translated in position, it results in an output which is the scaled and/or translated version of the output of the original input. Because of the difficulty in analyzing any nonlinear systems, theoretical work leading to the understanding of the behavior and properties of median filters has not been very successful. Properties of 2-D median filters are even less understood than the 1-D versions. Most of the theoretical work in understanding the deterministic properties of median filters concentrated on the study of the fixed-point sets mentioned in Section I. The fixed-point sets of a median filter are arrays unchanged or invariant under the operation of this particular filter. The edge preserving ability of median filters stems from the fact that the edges are unmodified or invariant to median filtering. Hence, one may state that the edges are fixed-point sets of median filters. The question arises as to which window shape and size will

preserve what type(s) of edges better? The same question may also be asked when it comes to preserve a certain desired object (or scene) in the image. This requires one to determine what type or filter window shape and size possesses what kind of fixed-point sets. There are results obtained for 1-D median filters,<sup>10, 13, 14, 15</sup> in which one has only to contend with window size. They cannot be extended to 2-D median filters because both shape and size are parameters. Nodes and Gallagher<sup>11</sup> attempted to extend some 1-D results to 2-D median filters via separable median filtering which operates on the image line-by-line and then follows it with column-by-column operation. It is not a true 2-D median filter, besides it can only imitate rectangular or square shape. Tyan<sup>10</sup> did a study on the fixed-point sets of 1-D median filters. He concluded that the necessary and sufficient condition for a sequence of numbers to be a fixed-point set of a median filter of window size  $2N+1$  is that every subsequence of size  $N+2$  must be monotonic or locally monotonic in  $N+2$ . This result does not apply for sequences of fast fluctuating binary numbers. He then attempted to extend his result to 2-D median filters and showed that it is sufficient to test the local monotonicity conditions of the pixel lines formed in all possible directions. For example, the sufficient condition for an array to be a fixed-point set of a  $3 \times 3$  window is that the array must be locally monotonic of length 3 for all the lines in all four directions which are  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ . The necessity of this condition has not been proven. It is seen that, since any linear subsequence for a single straight edge in any direction in an image is locally monotonic, the straight edge is preserved by a median filter of any symmetrical window as specified in Eq. (1). As far as the preservation of an object shape; (i.e., to determine whether this object shape is a fixed-point set of a median filter) is concerned, it depends on the relation between the object size and shape and the filter window size and shape. However, this quantitative relation is non-existent. In other words, there practically does not exist any fruitful result for determining fixed-point sets of 2-D median filters.

From the above discussion, the design of a median filter (i.e., the choice of filter window shape and size) depends on the noise to be reduced and the desired patterns or shapes to be preserved. Unless a quantitative description for the relation between the window shape and size and the patterns or object shapes can be found, there is no hope to design a median filter purely from theoretical point of view. Furthermore, any real-world image pattern can hardly meet the condition(s) of a fixed-point set. In practice, we do observe that some image patterns change very little after several passes through a median filter. We may say that these image patterns are approximate fixed-point sets for this particular median filter; however, the shape and size of this filter cannot be predetermined for a given image pattern.

It is also known that 1-D median filters are capable of changing a periodic sequence into a sequence of constant level when the window length is equal to the period. When the window length is not exactly equal to the period of the periodic sequence, it is likely that the periodic sequence is reduced to a ripple. Therefore, a 2-D median filter is capable of reducing any undesirable oscillating pattern in a picture while preserving edges.

Other means to characterize median filters have been proposed and studied but most of them are for 1-D filters as mentioned in Section I. Because of the non-linear nature, it is inappropriate to use impulse response, frequency response, and transfer function to characterize median filters. For example, Velleman<sup>32</sup> used a single-frequency signal of zero phase as an input to a 1-D median filter and obtained a transfer function for this single frequency but it cannot be generalized to other frequencies due to the nonlinearity of median filters. Kuhlman and Wise<sup>16</sup> analyzed the output autocorrelation function and the output power density spectrum of median filtered stationary sequences of independent data. The results of both Velleman and Kuhlman indicate a basically low-pass with some high-pass characteristics for 1-D median filters. They can hardly be generalized to 2-D filters. Examining the properties of median filters, we see that they



have no capability in restoring images degraded by blurs. It is seen that a median filter cannot be used alone as a preprocessor for both restoring blurred images and removing noise. Therefore, median filter must be used in conjunction with a filter which is properly designed for restoring blurred images. The inverse filter is chosen here for its relative simplicity in design and its perfect restoration of images to their original forms in the absence of noise.

The inverse filter is derived by using the least-squares criterion and on the basis that the observed image is modeled by

$$g(x,y) = h(x,y) \otimes f(x,y) + n(x,y) \quad (3)$$

where  $\otimes$  signifies the two-dimensional convolution;  $h(x,y)$  represents the PSF (Point Spread Function) which characterize the distortion (or blur);  $f(x,y)$ , the desired original image; and  $n(x,y)$ , the noise. As derived, the inverse filter expressed as transfer function is given as:

$$H_I(\omega_x, \omega_y) = \frac{1}{H(\omega_x, \omega_y)} \quad (4)$$

where  $H(\omega_x, \omega_y)$  is the OTF (Optical Transfer Function) which is the 2-D Fourier transform of the PSF  $h(x,y)$  in Eq. (3). To obtain the restored image, we multiply the Fourier transform of the observed image  $g(x,y)$  by  $H_I(\omega_x, \omega_y)$  and take the inverse Fourier transform of this product; i.e.,

$$\hat{f}(x,y) = \mathcal{F}^{-1}[H_I G] = \mathcal{F}^{-1}\left[\frac{G(\omega_x, \omega_y)}{H(\omega_x, \omega_y)}\right] \quad (5)$$

Using the operator matrix and vector notation we obtain

$$\hat{f} = H_I g = H^{-1}[Hf + n] = f + H^{-1}n \quad (6)$$

which is the original image plus the transformed noise. It is seen that, in the absence of noise and singularities (caused by the zero values of  $H$ ) of the inverse filter, the degraded image will be restored to the original form perfectly. In other words, the inverse filter concentrates all of its effort in regaining the original image  $f(x,y)$  but neglects (even amplifies)

the noise. In general,  $H(\omega_x, \omega_y)$  usually exhibit a low frequency "hump", hence  $1/H$  tends to boost the high-frequency portion of the observed image; i.e., the high-frequency components of the noise are usually accentuated in the restored image. By applying median filtering to this noisy restored image, it is expected that the median filter will remove the noise while not smearing the desired patterns in the image. On the other hand, one may apply median filtering first and followed by inverse filtering in the hope that the median filter will remove enough noise and not distort the blurred image so that the inverse filter would work more effectively in a less noisy environment. It is also possible to follow the above by another median filter to remove any accentuated noise due to inverse filtering.

The effectiveness of the combination of an inverse filter and a median filter as a preprocessor for either human processing or machine processing has never been explored and evaluated by other researchers. We attempt to do so in this project, though our approach is mainly empirical and exclusively for machine recognition.

#### IV. EXPERIMENT FOR PREPROCESSOR EVALUATION:

The experimental scheme is block-diagrammed as follows:

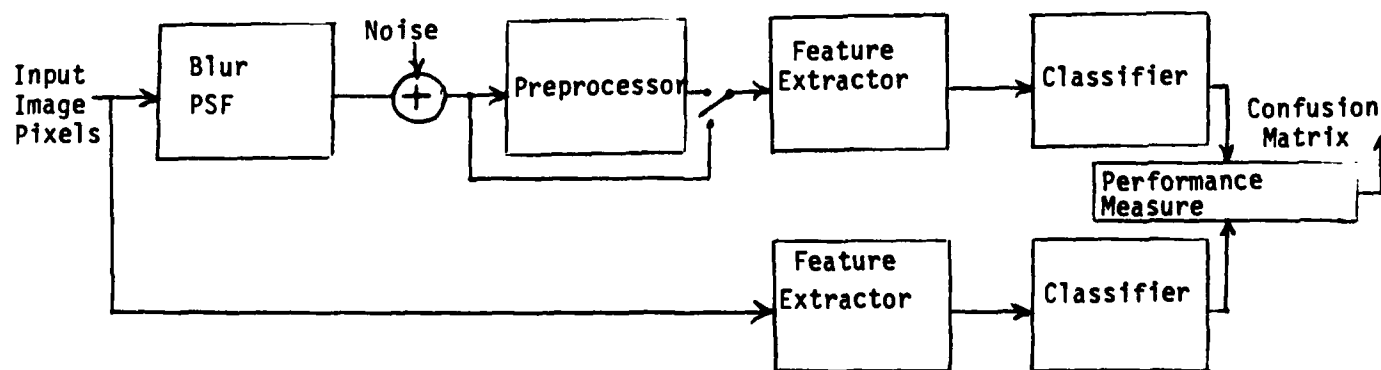


Fig 1 - Scheme For Experimentation

The preprocessor consists of either a median filter alone or a combination of inverse/median filters. Because of the limitation of time for this project, the input image data were taken from one picture only. The image selected is a section of an aerial photo of a site simulating a battleground scene. The image consists of a truck, its shadow and ground. This image was chosen since it contains simple truck and shadow shape with sharp edges and the scene in the image is simple enough for good machine classification. The image size is 128 x 128 pixels and each pixel value is represented by an 8-bit word. This input image, which will be referred as the original, and samples of degraded images which are a part of the data base for the preprocessor are shown in Fig 2. In order to evaluate filters under various operating conditions, a diverse group of degraded images was created by subjecting the original image to different types of degradation of various degrees of severity; i.e., several types of blur PSF's and additive noises at different levels were used in Fig. 1.

Three types of blurs, viz, camera-motion, defocused, and Gaussian blurs, were used because of their common occurrence. Two levels of severity for each type of blur were selected. The blurred images were created by convolving the original picture with the following blur PSF's<sup>33</sup> each at two levels of severity:

1. For the horizontal-camera-motion blur of moving length  $2d+1$  pixels, the PSF is expressed as

$$h(x,y) = \begin{cases} 0 & , y \neq 0, -\infty < x < \infty \\ \frac{1}{2d+1} & , y = 0, -d \leq x \leq d \end{cases} \quad (7)$$

$d = 5$  and  $10$  were chosen as low and high blurs, respectively

2. For the defocused blur by a defocused lens with circular aperture of radius  $R$ , the PSF is approximated by

$$h(x,y) = \begin{cases} 0 & , \sqrt{x^2+y^2} > R \\ \frac{1}{\pi R^2} & , \sqrt{x^2+y^2} \leq R \end{cases} \quad (8)$$

R=5 and 10 were used here.

3. For the Gaussian blur, the PFS is

$$h(x, y) = \frac{1}{2\pi\sigma_h^2} e^{-\frac{x^2+y^2}{2\sigma_h^2}} \quad (9)$$

$\sigma_h = 2.5$  and 5 were used.

These three different PSF's each with two levels of severity implemented as software and used to convolve with the original image to create six blurred images.

The original (unblurred) and the six blurred images were then subjected to further degradation by adding impulse noise (also known as salt-and-pepper noise) and Gaussian white noise. These noises were generated by the algorithms described below:

1. The impulse noise (salt-and-pepper or sp noise for short). This noise usually occurs when pixel values were wrongly received or decoded after transmission. In our experiment, it was generated by an algorithm<sup>9</sup> which simulates the above process. Each image pixel value  $g(x,y)$  at point  $(x,y)$  is replaced or not replaced by an erroneous value  $r(x,y)$  according to the rule:

$$g(x, y) = \begin{cases} r(x, y), & \text{with probability } P \\ g(x, y), & \text{with probability } 1-P \end{cases}$$

where  $r(x,y)$  is an independent random variable uniformly distributed in the interval 0,255 which is dictated by the 8-bit word length for each pixel value, and the probability  $P$  is independent of errors occurred at other points and the original pixel values.  $P = 0.1$  and  $0.2$  were used for evaluating median filter and inverse/median filters,  $P = 0.4$  (very severe) was used to generate a test image for filter window shape and size study, and  $P=0.5$  and  $0.6$  were also used to test the limits of median filtering in removing noise.

2. The Gaussian white noise. This noise was generated by adding, to the image, the uncorrelated random variable  $n(x,y)$  which is Gaussian distributed with the probability density function (p.d.f.),

$$p(n) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{n^2}{2\sigma_n^2}}$$

$\sigma_n = 6$  and 12 were used to generate two levels of noise.

It is seen that a data base consisting of 38 images were thus created. These include the original, 6 images degraded by blur only, 7 images degraded by noise only, and 24 degraded by both blur and noise in combinations of types of blurs and noises in different degrees of severity. Again, because of time limitation, only one degraded image was used for extensive testing in evaluating the effect of window shape and size on machine recognition. The image used was the original degraded only by impulse noise with  $P = 0.4$ . This severe noise was chosen based on the known property of median filter in removing impulse-type noise and the fact that moderate noise alone has little effect on machine recognition.<sup>27</sup>

Median filters of five window shapes (viz, square, cross, slanted cross, horizontal rectangle, and vertical rectangle) each with five different sizes were used as preprocessors. This resulted in 25 filtered images each of which was then machine classified. Their performances were compared. The median filter with the shape and size displaying the best performance was selected for use as the representative median filter for evaluating its effect on machine recognition under other image degradation conditions. That is, this median filter, and later on in conjunction with inverse filters, was used as the preprocessor in Fig. 1 on all of the images in the data base. The images with salt-and-pepper noise of  $P=0.5$  and 0.6 were used to evaluate the fall-off in the median filter performance.

In Fig's 2 and 3 examples of filtered images by this median filter and combination of this median filter and inverse filters are shown. All the median filtering performed here were single-pass filtering; i.e. the images were median-filtered only once. More passes may improve the performance.

A total of 127 images, which included all of the 38 images in the data base and the image preprocessed by various combination of median filtering and inverse/median filtering, were submitted for machine classification. The scene of each of those images was classified into three classes; viz., metal (truck), shadow, and ground (soil). Examples of machine classification are displayed in Fig's 4 and 5. The number of pixels classified into each class was then counted and compared to that of the true classifications to generate the conventional confusion matrices. The results of classification performed on the original image was taken as the reference (truth) based on the assumption that the machine can do no better on the degraded images than on the original. Based on the confusion matrix, the correct classification rate (CCR) for each class and the average correct classification rate (ACCR) were computed and used as the performance measures.

In the pattern recognition system, two features (average intensity over a  $3 \times 3$  window and contrast over a  $1 \times 5$  window were measured) and the condensed nearest neighbor classifier<sup>34</sup> was used. The features and the classifier were selected by experimentation on the original image and no attempt was made to find the best features and the best classifier for use on each of the images. Similarly, a set of training regions was determined by experimentation on the original and this set was used throughout the experiment so as to minimize the effect of training regions used. It is highly possible that one could better the classification performance on each image by choosing the best features, classifier, and training regions

for the particular image. Since our purpose is to evaluate and compare the preprocessors for facilitating machine recognition, the effect on classification performance due to other factors such as features, classifiers, training regions, etc. must be minimized. Therefore, in the strictest sense, our results pertain only to a specific set of experimental circumstances; i.e., a particular input image and a particular pattern recognition system. However, it is felt that the results obtained here may be generalized to other types of input images, pattern recognition systems, etc., as far as that they are applicable and thus shed light on the behavior of median filters and median/inverse filters as preprocessors for facilitating machine classification.

The preprocessor, feature extractor, etc are not isolated functional boxes as depicted in Fig 1; they interact to each other and are interrelated. Although we ignored the problem of feature extraction, classifier design, etc for the pattern recognition system in relation to the particular preprocessor used for machine recognition, the understanding of the interrelation between the preprocessor and the pattern recognition system is important for any practical application of machine recognition.

The experiment was carried out by using the following software packages implemented on DEC 11/70 and 11/34:

1. Median filters
2. Camera-motion blur, defocused blur, and pepper and salt noise generator.
3. "IRAMS" (Image Restoration and Manipulation Software)
4. "AFES" (Automatic Feature Extraction System).

The latter two were installed by PAR Technology Corp. for RADC. The former two were installed by us and designed to be, eventually, incorporated into IRAMS.

## V. RESULTS AND DISCUSSIONS

In the first part of the experiment, an image degraded by impulse noise (sp noise) of  $P = 0.4$  was used as the testing image and 25 median filters of 5 shapes and 5 sizes in combination were used as preprocessors shown in Fig. 1. The machine classification of all the filtered images and an unfiltered one resulted in 26 confusion matrices (as examples, some are shown in Table I). From which, the correct classification rate (CCR) for each class and the average CCR (ACCR) by class were computed and tabulated in Table II. It is seen that, in general, improvement in classification has been achieved by median filtering. However, one must be careful in choosing the window shape and size. For example, comparing Table I (b) and (d), we see that the  $11 \times 11$  slanted cross median filter improved the ground classification and practically no improvement for the shadow but wrongly classified many more truck pixels as ground and shadow. This is expected as discussed in Section III; since none of the arms of the slanted cross is perpendicular to the truck and shadow edges and the window size is too large for the truck and shadow, it smeared out the truck and shadow but filtered out enough noise to improve the ground classification. On the other hand, comparing Table I (b), (c) and (d), we see that the  $3 \times 3$  square window improved the truck and shadow classification by a large percentage but did not perform as well as the  $11 \times 11$  slanted cross in ground classification. From Table II, we observe that the  $3 \times 3$  square-window median filter performed the best based on ACCR and CCR for truck and shadow, i.e., it scored the highest ACCR and CCR for truck and shadow. Hence, the median filter of  $3 \times 3$  square window was selected as the representative preprocessor for use in the latter part of the experiment. To see the effect of window shape and size on classification in a more perspective way, ACCR and CCR versus the sizes of each window shape may be plotted using table II. Such plots of ACCR for the vertical rectangle and square are shown in Fig. 6. In the plot, the size  $1 \times 1$  signifies the unfiltered case. The plots for other window shapes have similar contour; they all peak at a particular size. It is interesting to note that the performance of the square window decays linearly as the size increases.



From Table II, we observe also that regardless of the window shape, the performance on soil classification increases as the size increases but it is just the opposite for truck and shadow classification. This is again as expected since the larger the window size the more effective in removing the noise but the more smearing of truck and shadow edges occurs. This agrees with the discussion in Section III.

A more extensive study on the effect due to salt-and-pepper noise levels (noise alone) on the classification performance of the  $3 \times 3$  square median filter was performed and results are shown in Table III. It shows that the filter has significant improvement on metal classification and more so as the noise gets heavier; the percentage of improvement increases from 7% at the noise level of  $P = 0.1$  to 24% at  $P = 0.6$ . It is also interesting to note that the CCR for each of the classes and the ACCR maintain a roughly constant level up to  $P = 0.4$  and drop off at  $P = 0.5$  and the percentage of improvement peaks about  $P = 0.4$  for ACCR. These results are plotted in Fig 7.

In general, we may conclude that the SP noise does not alter the performance until the level reaches a certain point. On the other hand, the improvement in performance is insignificant at low noise level and reaches a peak around the level  $P = 0.4$ . Even at very high noise level ( $P = 0.6$ ), there is still a classification improvement of 7.7% over the unfiltered case! Therefore, for images degraded by sp noise alone at moderate and higher levels it is most profitable to use median filter; at low noise level, the use of median filter will not improve the performance significantly.

For images degraded by Gaussian white noise alone at low and moderate levels, the classification performance was not affected substantially. This is seen from Table IV. It seems that the pattern recognition system can see through Gaussian white noise.<sup>27</sup> From Tables IV and V, we note that median filtering did improve the performance but not significantly. Hence, median filters are not recommended for use when images were degraded by Gaussian white noise alone especially at low and moderate levels.

For images degraded by blur and sp noise, the classification performance was enhanced by median filtering and more so at high noise level as shown in Tables IV and V. Comparing Tables IV and V, we see that median filtering has improved the performance for images degraded by all three types of blurs and sp noise and the improvement is more significant at higher noise level. From Table VI, we see that there is improvement in ACCR due to median filtering under all types of blurs and sp noise levels and the highest percentage is 22.5% improvement over the unfiltered case!

For images degraded by blur and Gaussian white noise, there is no discernable classification improvement due to median filtering as shown in Tables IV and V. From Table VI, we observe that there is even a slight regression in performance (as measured by ACCR) over the unfiltered case in all blur categories. Although median filtering does remove Gaussian white noise in terms of noise power,<sup>9</sup> it has little or no effect on classification performance. Hence, median filters should not be used to preprocess images degraded by Gaussian white noise and/or in combination with blurs for machine classification. This also agrees with the results reported in Ref. 27.

The study of the effect of median and inverse filtering in tandem on the classification performance had to be shortened due to the time limitation. However, we managed to obtain results for the case of defocused blur in combination with noises. These are shown in Tables VII and VIII. It is seen that, in general, there is significant improvement due to median/inverse filtering and especially so in cases of Gaussian white noise contamination; however, the most impressive improvement in performance is still in the sp noise contaminated cases; an improvement of about 30% over the unfiltered case was achieved. Therefore, for Gaussian-white-noise corrupted and blurred images, median/inverse filtering is helpful though median filtering alone is not. This shows again that median filtering does not enhance machine classification by removing white Gaussian noise since moderate Gaussian white noise affects classification performance insubstantially anyhow; however, median filtering does remove enough noise for inverse filter to work more effectively in the median/inverse filter arrangement. Through experimentation, we also determined that median filtering first and followed by inverse filtering is the best combination though the sequence med/inv/med has slight advantage however not worth the trouble.

To see how median/inverse filtering functioned, as an example, we shall discuss the case of low defocused blur ( $R=5$ ) with high salt-and-pepper noise ( $P = 0.2$ ). As seen in Fig's 4 and 5, and Table I (e), (f), (g), before filtering,  $1/3$  of the ground pixels were wrongly classified as metal; about  $1/5$  of metal pixels were wrongly classified as ground; and over  $1/2$  of the shadow pixels were wrongly classified as ground. After median filtering alone, the above confusion figures were reduced to:  $1/30$ ;  $1/10$ ; and  $1/4$ , respectively. This reduction in confusion is purely derived from removing noise without deblurring. With median/inverse filtering to first remove the noise and then restoring the image, the confusions were further reduced from  $1/3$  to  $1/100$ ;  $1/5$  to  $1/10$ ;  $1/2$  to  $1/50$ , respectively, though it increased the confusion in classifying ground pixels into shadow from negligible amount to  $1/30$ . Median/inverse filtering helped to separate ground from metal and shadow from ground. From this, it is not difficult to see how the improvement in classification performance due to median/inverse filtering as tabulated in Tables VII and VIII occurred? Although we did not evaluate the performance of median/inverse filtering for other blur cases, there is no reason to believe that the general conclusion about its effectiveness in enhancing machine classification would not apply for all types of blurs provided the inverse filter is properly designed.

In conclusion, we see that the median filtering offers the most effective preprocessing for enhancing machine recognition in cases when images are contaminated by impulse-type noise, and it is also effective in cases when images are degraded by blurs and impulse-type noise. Median/inverse filtering will improve the latter cases even further. Median filtering alone is not effective in facilitating machine recognition when images are corrupted by Gaussian white noise and/or in combination with blurs; however, median/inverse filtering does improve the performance for cases of combinations of Gaussian white noise and blurs.

## VI. RECOMMENDATIONS:

Because of time limitations, we could not perform an exhaustive performance evaluation of the median filter window shape and size when operating against various degradation conditions. Instead, we optimized the shape/size parameters of the median window for one specific degradation condition; i.e., severe salt and pepper noise ( $p = 0.4$ ). Hence, our first recommendation would be:

(1) Evaluate the performance of median filters of various shapes and sizes operating under various degradation conditions so as to optimize window shape/size for each degradation condition.

From the results shown in this study, 2-D median filter can serve effectively as a preprocessor for automatic pattern recognition systems. Our second recommendation would be to exploit its simplicity and relative ease in implementation to realize its real-time processing potential; to be more specific, this recommendation is stated as:

(2) Design and implement a microprocessor-based median filter or, if possible, a VLSI (or VHSI) median-filter chip aimed at real-time filtering.

Again, we did not have time to optimize the design of inverse filters in our study of the effect of median-inverse filtering. Though our results show improvement in performance over median filtered alone and unfiltered cases, we believe the performance could be enhanced even more if the inverse filters were better designed. Hence, the third recommendation is:

(3) Extend our study on the median-inverse filtering in tandem by optimizing the design of inverse filters.

Other recommendations for further research include:

(4) Study and evaluate other variant median filters or generalized median filters, some of which were discussed in Section III.

(5) Study and evaluate recursive median filtering which may speed up processing.

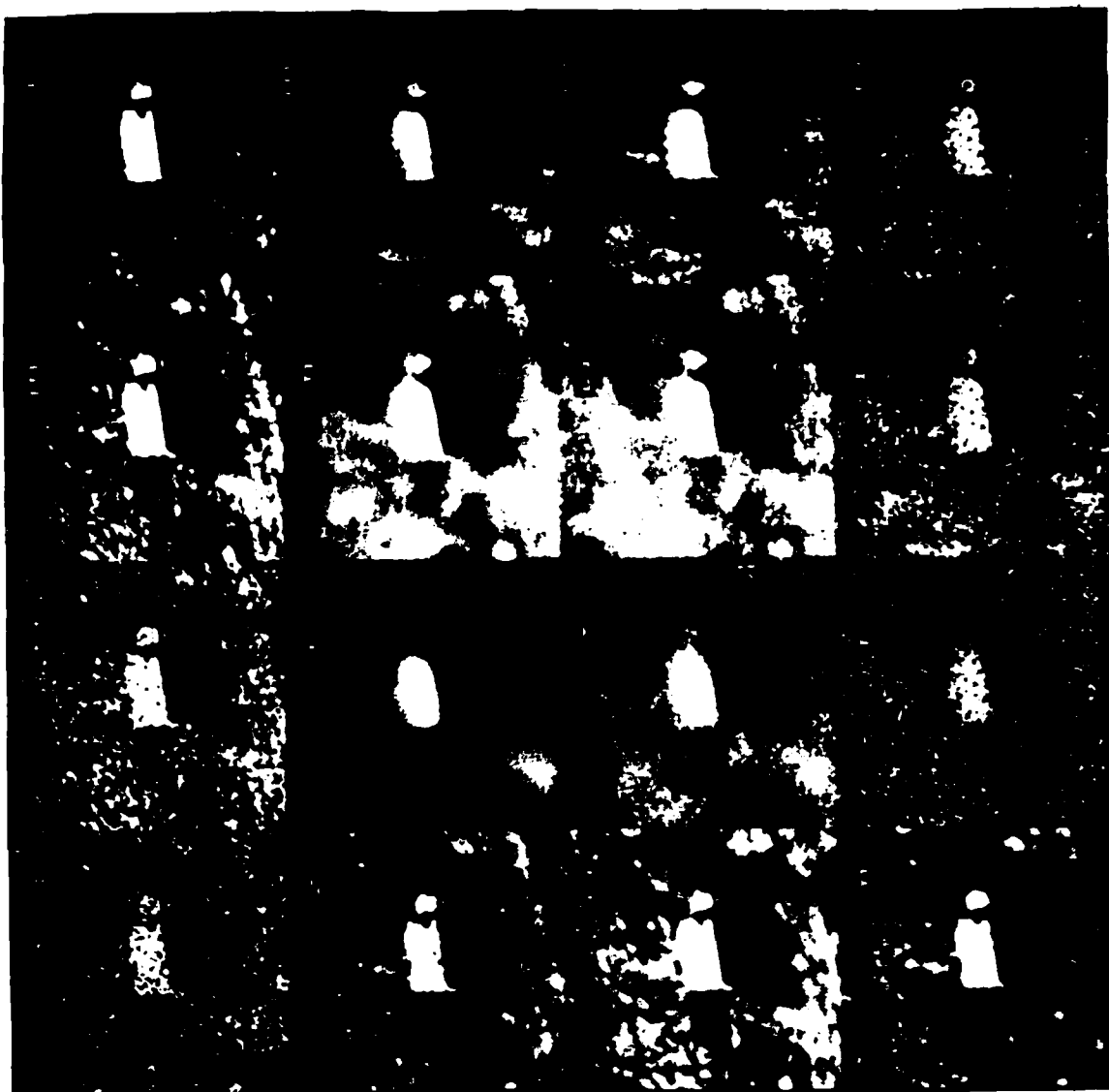
(6) Evaluate median filters not just for pattern recognition (as it was done here) but for scene analysis or image understanding by using AI rules; i.e., going a step beyond pixel classification as it was performed here.

(7) Evaluate median filtering for its effect in removing scanner noise, noises having oscillating or periodic structure, noise caused by missing data, etc.

(8) Study and understand how image filters and/or enhancement techniques interrelate to feature extractor, classifier, selection of training regions if supervised learning is used, and the selection of AI rules with the aim to develop machine "recognizability function" or machine "interpretability function" (versus the human "visibility function"<sup>35</sup>).

(9) Because of the ill-conditioned nature of the image restoration problem, there are infinite variety of solutions (filters). The filters (solutions) have been customarily designed (selected) according to different sets of mathematical criteria, e.g., the minimum-mean-square error criterion, based on various assumptions of the image models. It would be more desirable to design filters (i.e., to seek solutions) to optimize the machine "recognizability function" or "interpretability function" mentioned above. Hence, this recommendation involves: Study the feasibility of designing image restoration filters to optimize the machine "recognizability function" instead of mathematical criteria or human "visibility function."<sup>36</sup>

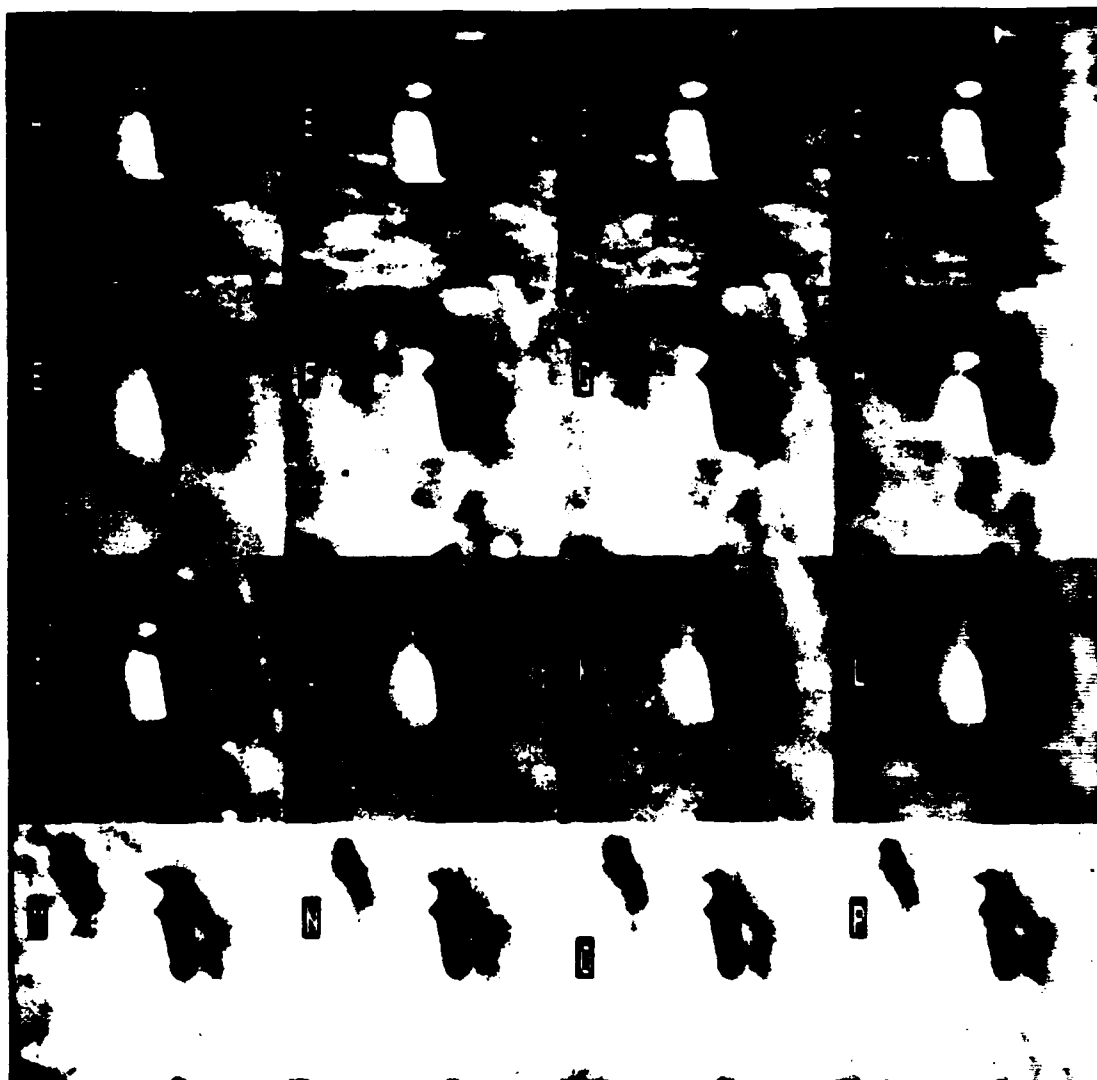
(10) Design and evaluate adaptive filters in general or adaptive median filters in particular, to adapt to the changing conditions of the pattern recognition system (or the AI scene analyzer) as specified by the recognizability function or interpretability function instead of adapting to human "visibility function."<sup>37</sup>



- |                                            |                                                     |
|--------------------------------------------|-----------------------------------------------------|
| (A) Original                               | (I) SPN ( $P = 0.2$ )                               |
| (B) MB ( $d=5$ )                           | (J) GB ( $\sigma_h = 5$ )                           |
| (C) MB ( $d=5$ ) + GWN ( $\sigma_n = 12$ ) | (K) GB ( $\sigma_h = 5$ ) + GWN ( $\sigma_n = 12$ ) |
| (D) MB ( $d=5$ ) + SPN ( $P = 0.2$ )       | (L) GB ( $\sigma_h = 5$ ) + SPN ( $P = 0.2$ )       |
| (E) GWN ( $\sigma_n = 12$ )                | (M) SPN ( $P = 0.4$ )                               |
| (F) DB ( $R=5$ )                           | (N) MF SPN ( $P = 0.4$ )                            |
| (G) DB ( $R=5$ ) + GWN ( $\sigma_n = 12$ ) | (O) MF GWN ( $\sigma_n = 12$ )                      |
| (H) DB ( $R=5$ ) + SPN ( $P=0.2$ )         | (P) MF SPN ( $P = 0.2$ )                            |

MB: Motion Blur    DB: Defocused Blur    GB: Gaussian Blur  
 SPN: Salt-and-Pepper Noise    GWN: Gaussian White Noise  
 MF: Median Filtered    M/I F: Median/Inverse Filtered

Fig. 2 Portion of Data Base



- |                                           |                                                  |
|-------------------------------------------|--------------------------------------------------|
| (A) MF MB ( $D=10$ )+SPN ( $P=0.2$ )      | (I) MF GB ( $\sigma_n = 2.5$ ) + SPN ( $P=0.2$ ) |
| (B) MF MB ( $d=5$ )                       | (J) MF GB ( $\sigma_n = 5$ )                     |
| (C) MF MB ( $d=5$ )+GWN ( $\sigma_n=12$ ) | (K) MF GB ( $\sigma_n=5$ )+GWN ( $\sigma_n=12$ ) |
| (D) MF MB ( $d=5$ )+SPN ( $P=0.2$ )       | (L) MF GB ( $\sigma_n=5$ )+SPN ( $P=0.2$ )       |
| (E) MF DB ( $R=10$ )+SPN ( $P=0.2$ )      | (M) M/I F DB ( $R=5$ )+SPN ( $P=0.1$ )           |
| (F) MF DB ( $R=5$ )                       | (N) M/I F DB ( $R=5$ )+GWN ( $\sigma_n = 5$ )    |
| (G) MF DB ( $R=5$ )+GWN ( $\sigma_n=12$ ) | (O) M/I F DB ( $R=5$ )+SPN ( $P=0.1$ )           |
| (H) MF DB ( $R=5$ )+SPN ( $P=0.2$ )       | (P) M/I F DB ( $R=5$ )+SPN ( $P=0.2$ )           |

Legend: Same as Fig 2

Fig. 3 Examples of Filtered Images

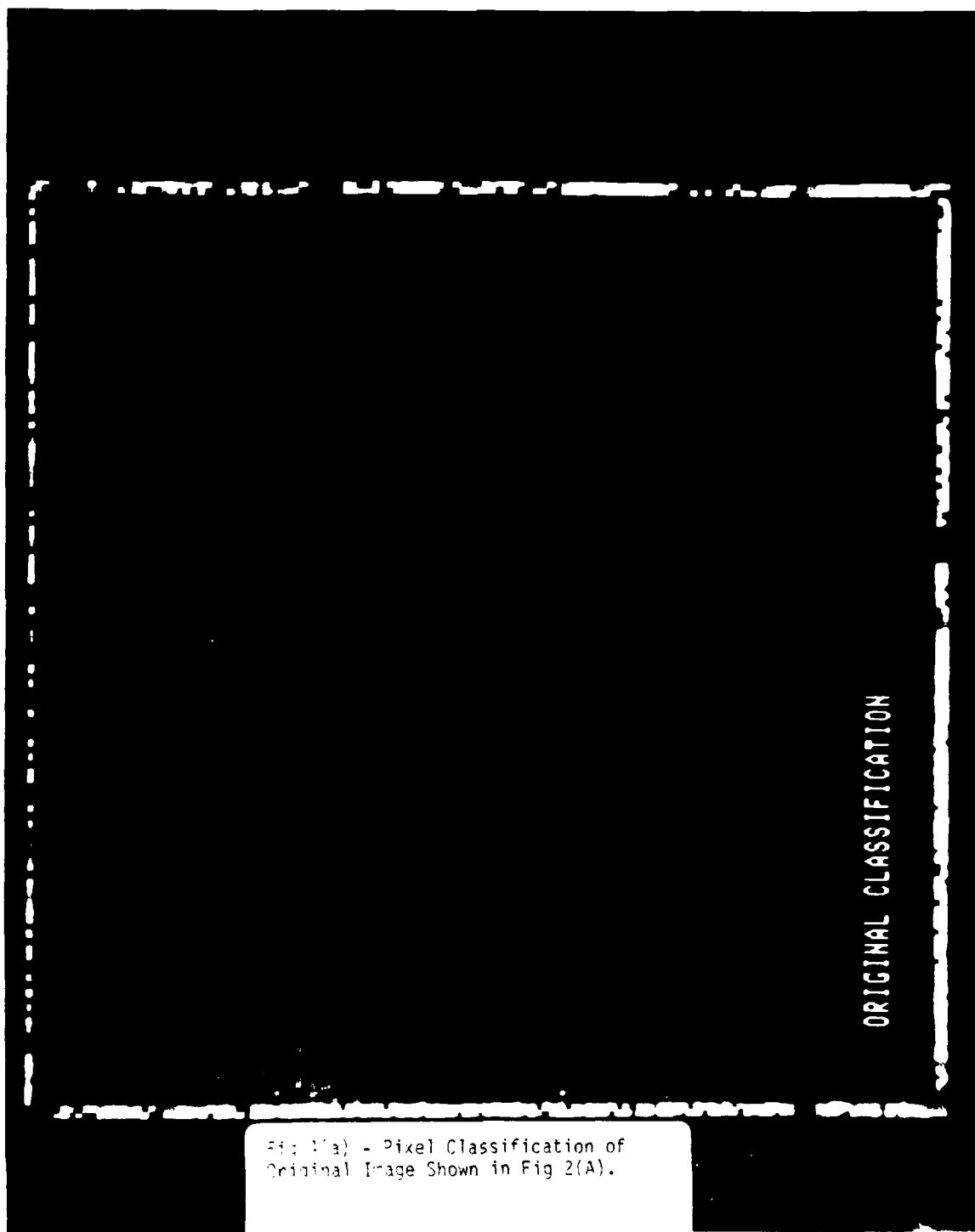


Fig. 1(a) - Pixel Classification of  
Original Image Shown in Fig 2(A).



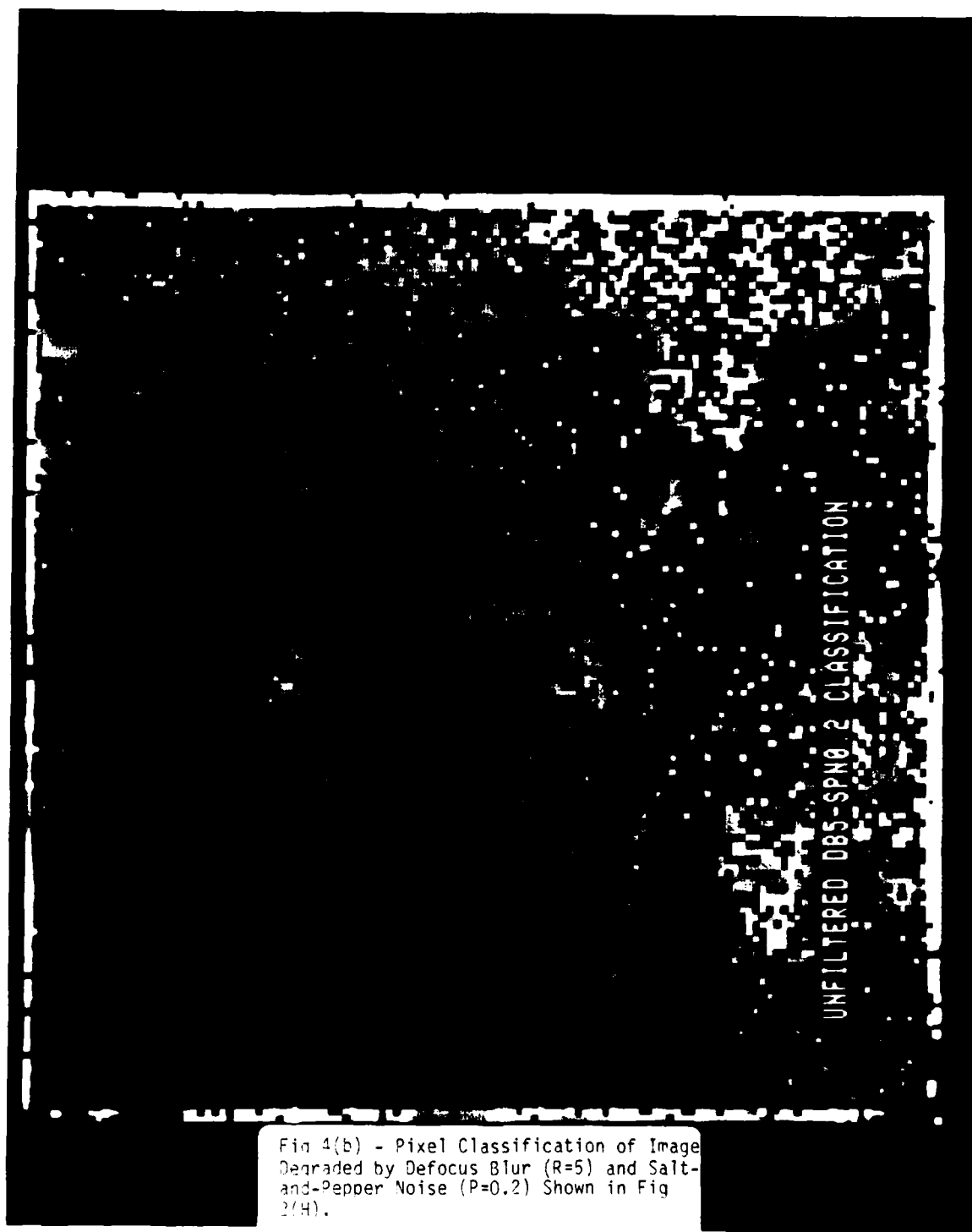
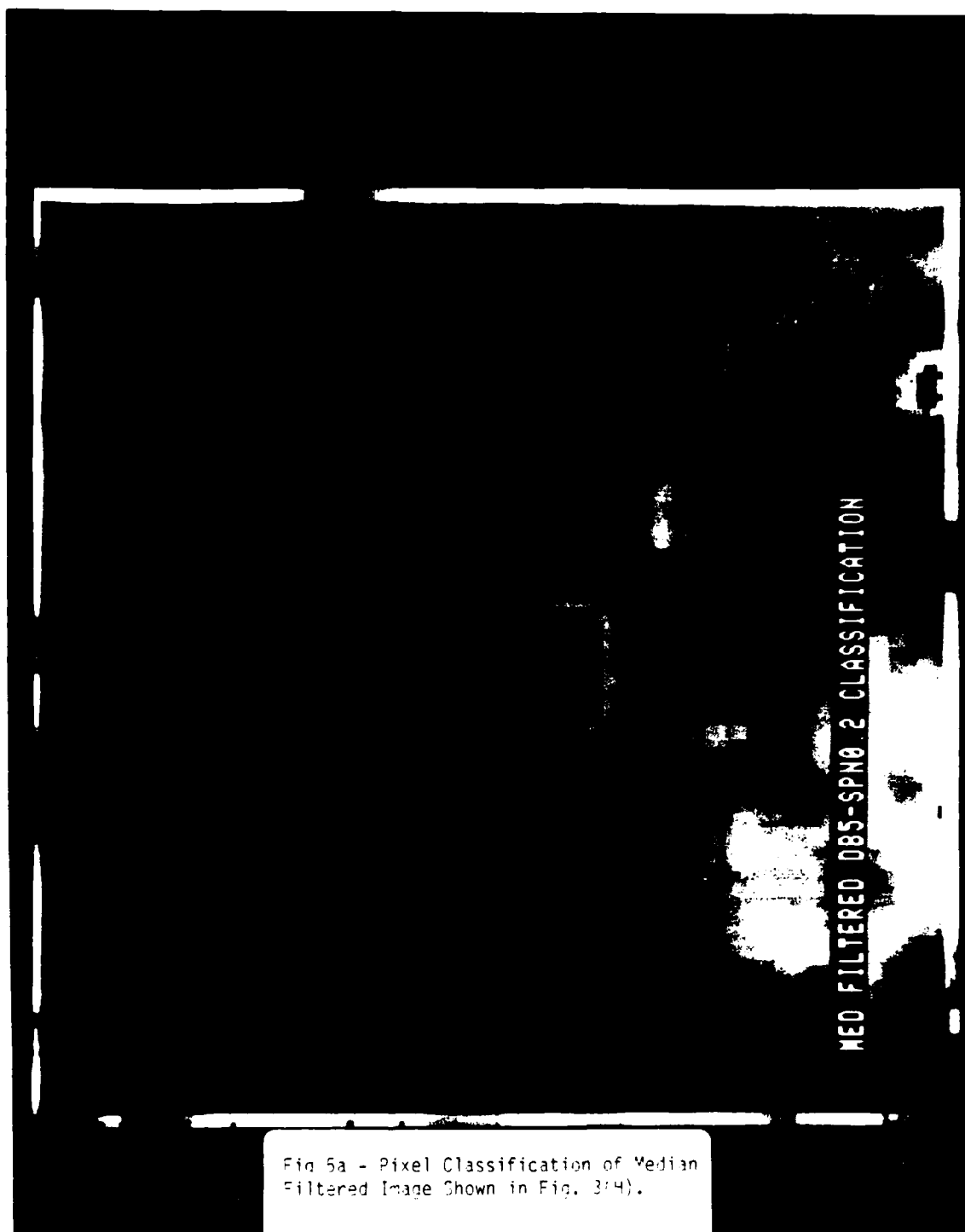
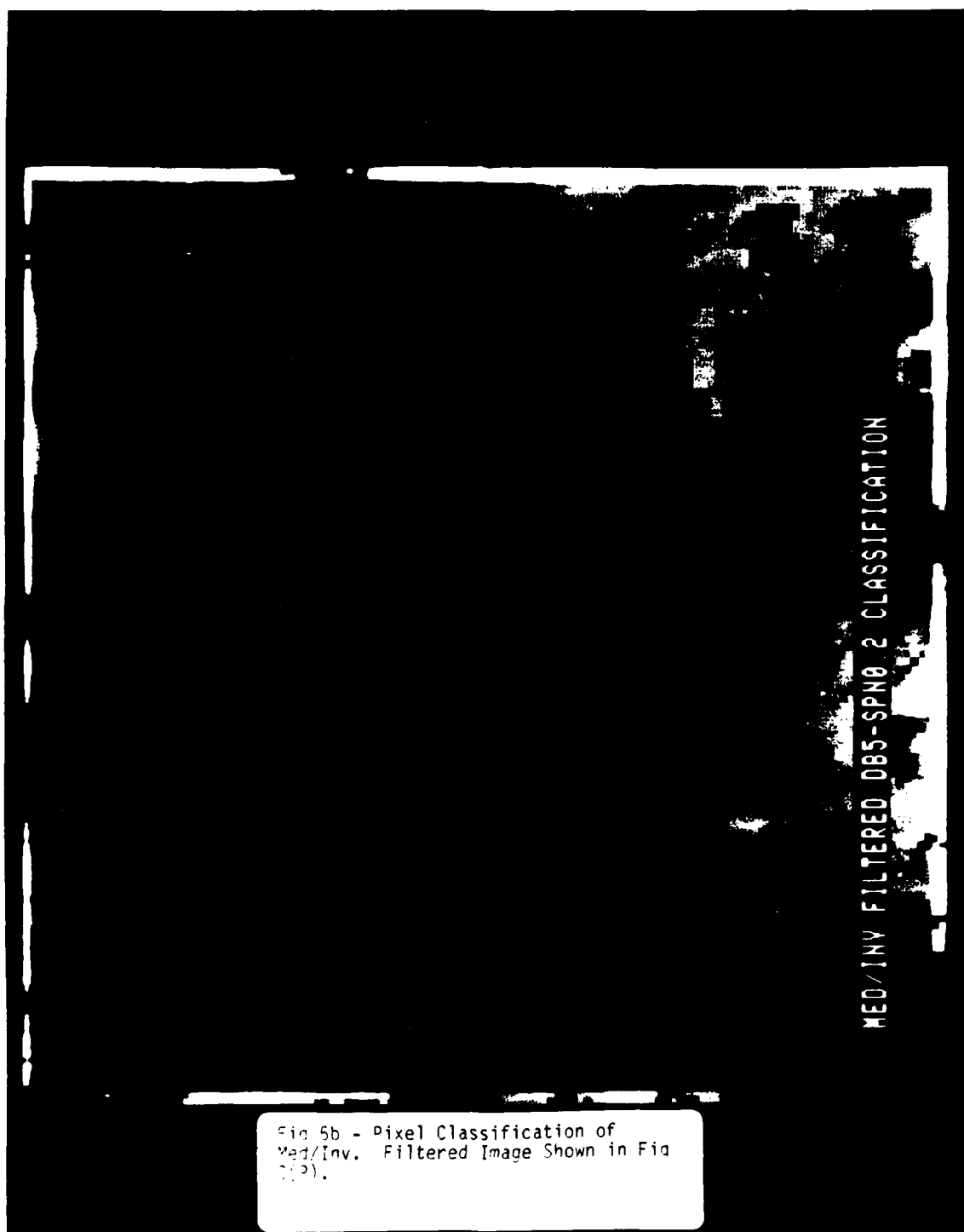


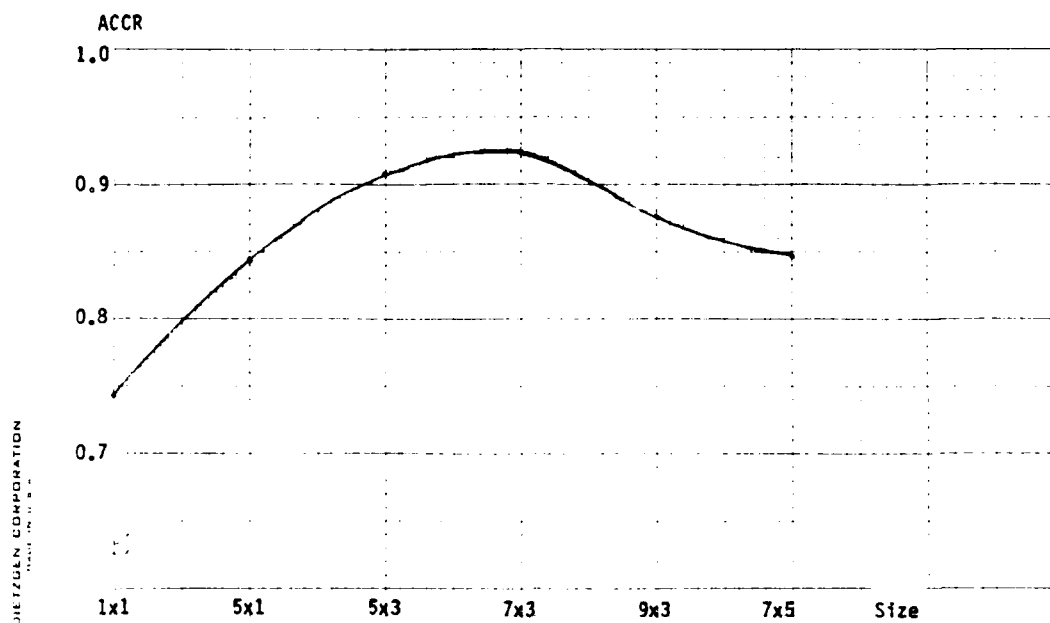
Fig 4(b) - Pixel Classification of Image Degraded by Defocus Blur ( $R=5$ ) and Salt-and-Pepper Noise ( $P=0.2$ ) Shown in Fig 2(H).



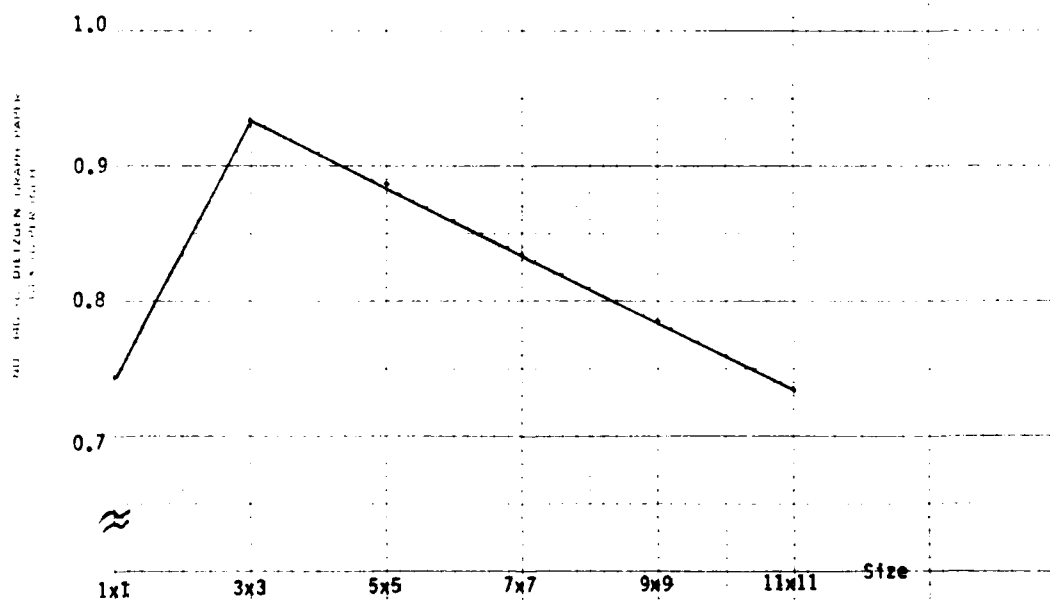
MED FILTERED DBS-SPN0.2 CLASSIFICATION

Fig 5a - Pixel Classification of Median Filtered Image Shown in Fig. 3(H).



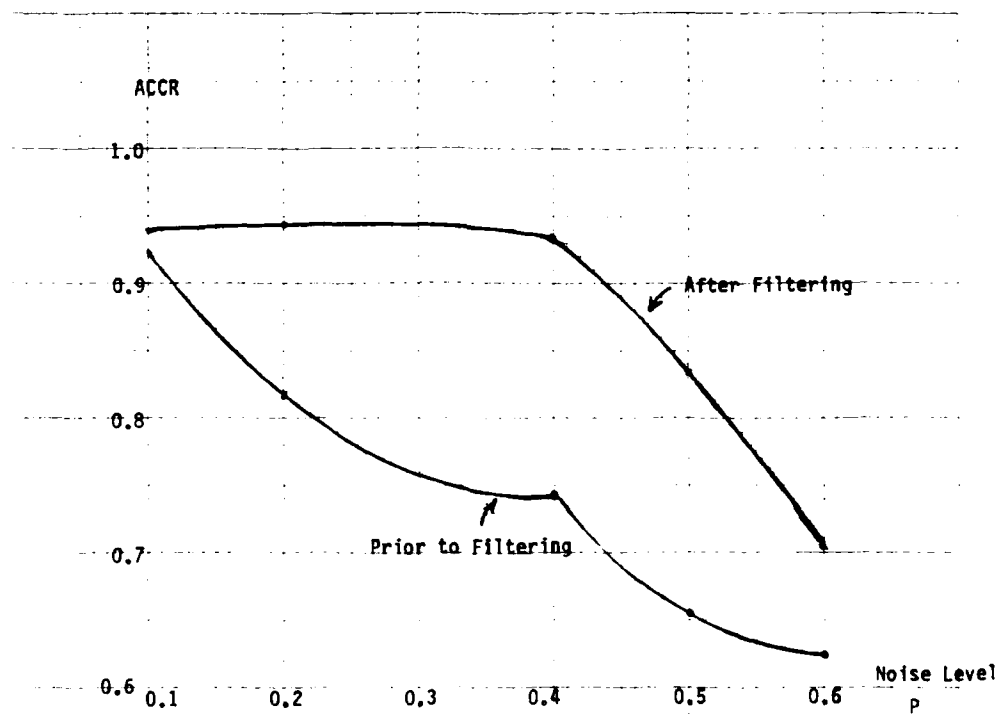


(a) Vertical Rectangular Median Filter



(b) Square Median Filter

FIG 6. Average Correct Classification Rate vs Window Size



(a) Average Correct Classification Rate

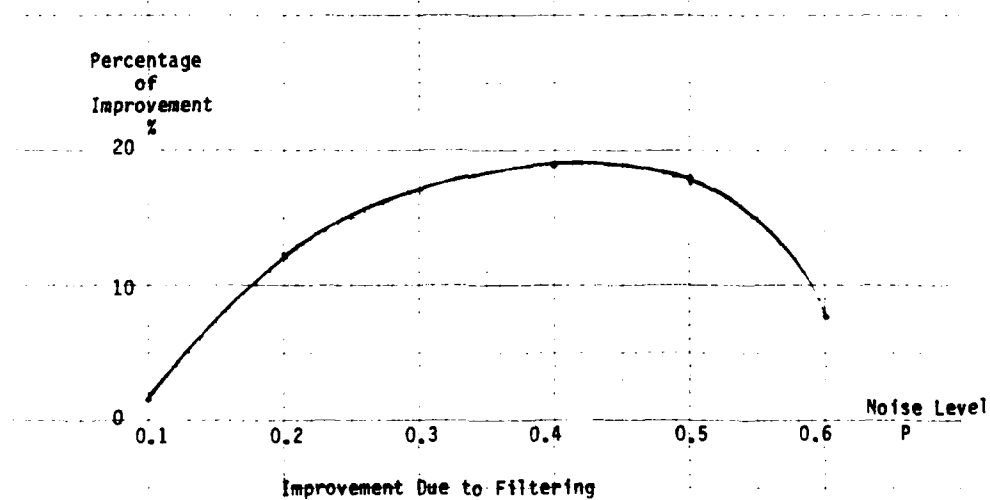


Fig. 7 - Effect of SP Noise Level on Classification Performance

TABLE I - Examples of Confusion Matrices

	M	S	G
M	602	0	0
S	0	482	0
G	0	0	14416

(a) Original Image

	M	S	G
M	414	0	188
S	0	332	150
G	1911	159	12346

(b) sp noise with  $P=0.4$  without filtering.

	M	S	G
M	567	0	35
S	0	420	62
G	122	36	14258

(c) 3 x 3 square med. filtering.  
on image with sp noise ( $P=0.4$ )

	M	S	G
M	346	8	248
S	0	336	146
G	3	77	14336

(d) 11x11 slanted cross med. filtering  
on image with sp noise ( $P=0.4$ )

	M	S	G
M	486	0	116
S	0	217	265
G	4590	44	9782

(e) Image degraded by defocused  
blur ( $R=5$ ) and sp noise ( $P=0.2$ )

	M	S	G
M	549	0	53
S	0	355	127
G	482	64	13870

(f) 3 x 3 Med. filtering of image  
degraded by defocused blur ( $R=5$ ) and SPN  
( $P=0.2$ )

	M	S	G
M	530	0	72
S	0	474	8
G	155	476	13785

(g) Med/Inverse filtering of Image  
degraded by defocused blur ( $R=5$ ) and  
SPN ( $P=0.3$ )

M: Metal (Truck)

S: Shadow

G: Ground (soil)

TABLE II - CCR and ACCR for Median Filters of Various Shapes  
and Sizes on an Image with Salt-and-Pepper Noise at the level of  $P = 0.4$

(a) CCR for Metal (Truck)

Shape \ Size	A	B	C	D	E
SQ	.9419	.8455	.8173	.6611	.6545
CR	.8771	.8738	.8787	.8505	.8605
SC	.8538	.8007	.6512	.6910	.5748
HR	.8239	.9369	.8073	.8721	.7625
VR	.8970	.8621	.9402	.7691	.8239

Unfiltered: .6877

(b) CCR for Shadow

Shape \ Size	A	B	C	D	E
SQ	.8714	.8216	.6763	.7012	.5498
CR	.8112	.8091	.8465	.7303	.7324
SC	.8444	.8320	.5851	.6162	.6971
HR	.8257	.8257	.6515	.6971	.7158
VR	.7822	.8693	.8382	.8651	.7241

Unfiltered: .6888

(c) CCR for Soil

Shape \ Size	A	B	C	D	E
SQ	.9890	.9974	.9985	.9983	.9993
CR	.8880	.9968	.9953	.9946	.9967
SC	.9786	.9905	.9968	.9968	.9954
HR	.9550	.9857	.9986	.9877	.9987
VR	.8502	.9960	.9898	.9977	.9992

Unfiltered: .8564

TABLE II (CONT)

(d) ACCR

Shape	Size				
	A	B	C	D	E
SQ	.9341	.8882	.8307	.7869	.7345
CR	.8588	.8932	.9068	.8585	.8632
SC	.8923	.8744	.7443	.7680	.7554
HR	.8682	.9161	.8191	.8523	.8256
VR	.8431	.9092	.9227	.8773	.8491

Unfiltered: .7443

Legend:

SQ: Square; CR: Cross; SC: Slanted Cross;

HR: Horizontal Rectangle; VR: Vertical Rectangle

	A	B	C	D	E
SQ	3x3	5x5	7x7	9x9	11x11
CR	3x3	5x5	7x7	9x9	11x11
SC	3x3	5x5	7x7	9x9	11x11
HR	1x5	3x5	3x7	3x9	5x7
VR	5x1	5x3	7x3	9x3	7x5



TABLE III CCR and ACCR for the 3 x 3 Median Filter  
on Images with Salt-and-Pepper Noise at Various Levels

Noise Levels Classes		0.1	0.2	0.4	0.5	0.6
M	PF	.9302	.8422	.6877	.6761	.5797
	AF	1.0000	.9568	.9419	.8538	.8156
S	PF	.8963	.7178	.6888	.5021	.5768
	AF	.9544	.8734	.8714	.7593	.8008
G	PF	.9437	.8947	.8564	.7867	.7168
	AF	.8638	.9977	.9890	.8928	.4913
Average	PF	.9234	.8182	.7443	.6550	.6254
	AF	.9394	.9427	.9341	.8353	.7026
	%	1.6	12.45	18.98	18.03	7.72

M: Metal (Truck)      S: Shadow      G: Ground (soil)

PF: Prior to Filtering    AF: After Filtering

%: Percentage of Improvement due to Filtering

TABLE IV CCR and ACCR for Unfiltered Images

Noise Blur	No Noise	SPN		GWN	
		0.1 <sup>(P)</sup>	0.2	6 <sup>(<math>\sigma_n</math>)</sup>	12
No Blur	1.0000	.9302	.8422	1.000	1.000
GB ( $\sigma_n$ )	2.5	.8688	.8439	.7475	.7957
	5	.7442	.7292	.7409	.7575
DB (R)	5	.8173	.7973	.8073	.8920
	10	.6545	.7292	.6894	.7525
MB (d)	5	.9103	.8638	.7907	.9120
	10	.8472	.8837	.9186	.7475

(a) CCR for Metal (Truck)

Noise Blur	No Noise	SPN		GWN	
		0.1 <sup>(P)</sup>	0.2	6 <sup>(<math>\sigma_n</math>)</sup>	12
No Blur	1.0000	.8963	.7178	.7946	.8071
GB ( $\sigma_n$ )	2.5	.7261	.6639	.6432	.8174
	5	.8237	.7842	.5145	.7095
DB (R)	5	.7324	.7593	.4502	.7241
	10	.7199	.5228	.6494	.7510
MB (d)	5	.7822	.7178	.6763	.7261
	10	.9108	.5311	.6058	.9066

(b) CCR for shadow

Noise Blur	No Noise	SPN		GWN	
		0.1 <sup>(P)</sup>	0.2	6 <sup>(<math>\sigma_n</math>)</sup>	12
No Blur	1.0000	.9437	.8947	.9797	.9900
GB ( $\sigma_n$ )	2.5	.9954	.9738	.7538	.9942
	5	.9542	.8678	.6099	.9468
DB (R)	5	.9911	.9718	.6786	.9778
	10	.9711	.7611	.6082	.8308
MB (d)	5	.9650	.8236	.6556	.9733
	10	.7669	.4821	.3563	.9007

(c) CCR for Ground (soil)

Noise Blur	No Noise	SPN		GWN	
		0.1 <sup>(P)</sup>	0.2	6 <sup>(<math>\sigma_n</math>)</sup>	12
No Blur	1.0000	.9234	.8182	.9248	.9324
GB ( $\sigma_n$ )	2.5	.8634	.8272	.7148	.8691
	5	.8407	.7938	.6218	.8046
DB (R)	5	.8469	.8428	.6454	.8646
	10	.7818	.6711	.6490	.7781
MB (d)	5	.8858	.8017	.7075	.8705
	10	.8416	.6323	.6269	.8516

(d) Average CCR by class

GB: Gaussian Blur

DB: Defocused Blur

MB: Motion Blur

SPN: Salt-and-Pepper Noise

GWN: Gaussian White Noise

TABLE V CCR and ACCR for Images Preprocessed  
by Median Filtering

Noise Blur	No Noise	SPN		GWN	
		(P)		(σ <sub>n</sub> )	
		0.1	0.2	6	12
No Blur	N.A.	1.0000	.9568	.9751	.9718
GB (σ <sub>n</sub> )	2.5	.9336	.8140	.7425	.9850
	5	.7442	.8090	.8289	.7010
DB (R)	5	.7326	.8704	.9120	.8439
	10	.8389	.7110	.7243	.7259
MB (d)	5	.9186	.7243	.8987	.9452
	10	.6694	.8007	.8588	.7741

(a) CCR for Metal (Truck)

Noise Blur	No Noise	SPN		GWN	
		(P)		(σ <sub>n</sub> )	
		0.1	0.2	6	12
No Blur	N.A.	.9544	.8734	.8942	.8506
GB (σ <sub>n</sub> )	2.5	.7552	.7469	.7842	.6846
	5	.8237	.7344	.7739	.7324
DB (R)	5	.7324	.6950	.7365	.6743
	10	.7925	.7905	.8444	.6888
MB (d)	5	.8071	.7490	.8610	.7863
	10	.8714	.6722	.6701	.9253

(b) CCR for Shadow

Noise Blur	No Noise	SPN		GWN	
		(P)		(σ <sub>n</sub> )	
		0.1	0.1	6	12
No Blur	N.A.	.8638	.9977	.9994	.9993
GB (σ <sub>n</sub> )	2.5	.9834	.9956	.9950	.4870
	5	.9544	.9006	.9539	.9800
DB (R)	5	.9947	.9822	.9621	.9879
	10	.5325	.9351	.8876	.9160
MB (d)	5	.9535	.9801	.9652	.8285
	10	.9440	.8899	.6598	.8759

(c) CCR for Ground (Soil)

Noise Blur	No Noise	SPN		GWN	
		(P)		(σ <sub>n</sub> )	
		0.1	0.2	6	12
No Blur	N.A.	.9394	.9427	.9562	.9406
GB (σ <sub>n</sub> )	2.5	.8907	.8521	.8694	.7189
	5	.8407	.8147	.8234	.8044
DB (R)	5	.8199	.8492	.8702	.8353
	10	.7213	.8122	.8188	.7769
MB (d)	5	.8930	.8709	.9083	.8533
	10	.8283	.7876	.7296	.8584

(d) Average CCR by Class.

GB: Gaussian Blur      DB: Defocused Blur      MB: Motion Blur  
SPN: Salt-and-Pepper Noise      GWN: Gaussian White Noise

TABLE VI

Percentage of Improvement in ACCR  
by Median Filtering

Noise Blur		No	SPN		GWN	
		Noise	0.1	(P) 0.2	6	(σ <sub>n</sub> ) 12
No Blur			1.60	12.45	3.14	0.82
GB	2.5	2.72	2.49	15.45	-15.02	0.79
	5	0	2.09	20.16	-0.02	-0.48
DB (R)	5	-2.70	0.64	22.48	-2.93	-0.19
	10	-6.05	14.11	16.98	-0.12	-0.27
MB (d)	5	0.72	6.92	20.08	-1.72	-4.39
	10	-1.33	15.53	10.27	0.68	-0.68

**TABLE VII CCR and ACCR for Noisy Defocus Blurred  
Images Preprocessed by Median/Inverse Filtering**

Class	Noise Blur Levels	No Noise	SPN		GWN	
			(P)		$(\sigma_n)$	
			0.1	0.2	6	12
M	5	.9385	.8887	.8804	.9086	.9302
	10	.6645	.6827	.7193	.7093	.6944
S	5	.8714	.7780	.9834	.7697	.6494
	10	.9896	.9917	.7739	.9730	.7759
G	5	.9587	.9837	.9562	.9779	.9690
	10	.8941	.8829	.9134	.8903	.9560
Ave	5	.9229	.8835	.9400	.8854	.8495
	10	.8494	.8524	.8022	.8576	.8088

**TABLE VIII. Percentage of Improvement in ACCR by  
Median/Inverse Filtering**

Comparison	Noise Blur Level (R)	No Noise	SPN		GWN	
			(P)		$(\sigma_n)$	
Base % Over Unfiltered	5	7.60	4.07	29.46	2.08	1.26
	10	6.76	18.13	15.32	7.95	13.22
% Over Med Filtered	5	10.30	3.43	6.98	5.01	1.45
	10	12.81	4.02	1.66	8.07	13.49

#### REFERENCES

1. J. W. Tukey, Exploratory Data Analysis, Reading, MA: Addison-Wesley, 1977 (1st Ed. 1971).
2. L. R. Rabiner, M.R. Sambur, and C.E. Schmidt, "Applications of a Nonlinear Smoothing Algorithm to Speech Processing", IEEE Trans. ASSP-23, pp. 552-557, Dec 1975.
3. N.S. Jayant, "Average and Median Based Smoothing Techniques for Improving Digital Speech Quality in the Presence of Transmission Errors", IEEE Trans, COM-24, pp. 1043-1045, Sept 1976.
4. R. Steele and D.J. Goodman, "Detection and Selective Smoothing of Transmission Errors in Linear PCM", BSTJ, Vol 56, pp. 300-409, March 1977.
5. W. K. Pratt, Digital Image Processing, New York, N.Y.: Wiley Interscience, 1978, pp 330-333.
6. B. R. Frieden, "A New Restoring Algorithm for the Preferential Enhancement of Edge Gradients", J. Opt. SOC Am., Vol 66, pp. 280-283, 1976.
7. T. S. Huang (Ed.), Two-Dimensional Digital Signal Processing II, Berlin, W. Germany: Springer-Verlag, 1981, pp. 3-6.
8. G. W. Wecksung and K. Campbell, "Digital Image Processing at EG&G", IEEE Computer Vol 7, No. 5, pp. 63-71, May 1974.
9. B. I. Justusson, "Median Filtering: Statistical Properties", Two-Dimensional Digital Signal Processing II, pp. 161-196 (Editor: T. S. Huang), Berlin, W. Germany: Springer-Verlag, 1981.
10. S. G. Tyan, "Median Filtering: Deterministic Properties", Two-Dimensional Digital Signal Processing II, pp. 197-217 (Editor: T. S. Huang), Berlin, W. Germany: Springer-Verlag, 1981.
11. T. A. Nodes and N. C. Gallagher, Jr., "Two-Dimensional Root Structures and Convergence Properties of the Separable Median Filter", IEEE Trans. ASSP-31, pp. 1360-1365.
12. E. Ataman, V. K. Aatre, and K. M. Wong, "Some statistical properties of Median Filters" (Correp.), IEEE Trans. ASSP-29, pp. 1073-1075, Oct 1981.
13. N. C. Gallagher, Jr. and G. L. Wise, "A Theoretical Analysis of the Properties of Median Filters" IEEE Trans. ASSP-29, pp. 1136-1141, Dec 1981.
14. T. A. Nodes and N. C. Gallagher, Jr., "Median Filters: Some Modifications and Their Properties," IEEE Trans. ASSP-30, pp. 739-745, Oct 1982.

15. G. R. Arce and N. C. Gallagher, Jr., "State Description of the Root-Signal Set of Median Filters", IEEE Trans. ASSP-30, pp. 894-902, Dec 1982.
16. F. Kuhlmann and G. L. Wise, "On Second Moment Properties of Median Filtered Independent Data" IEEE Trans. COM-29 pp. 1374-1379, Sept 1981.
17. T. S. Huang, G. J. Yang, and G. Y. Tang, "A Fast Two-Dimensional Median Filtering Algorithm", IEEE Trans. ASSP-27, pp. 13-18, Feb 1979.
18. E. Ataman, V. K. Aatre, and K. M. Wong, "A Fast-Method for Real-Time Median Filtering" IEEE Trans. ASSP-28, pp. 415-421, Aug 1980.
19. G. Garibotto and L. Lambarelli, "Fast On-line Implementation of Two-Dimensional Filtering", Electronic Letter, Vol 15, pp. 24-25, Jan 1979.
20. G. P. Wolfe and J. L. Mannos, "A Fast Median Filter Implementation", Proc. SPIE Seminar on Image Processing, Sept 1979, San Diego, CA.
21. W. L. Eversole, D. J. Mayer, F.B. Frazee, and T. F. Cheek, Jr., "Investigation of VLSI Technologies for Image Processing" Proc. Image Understanding Workshop, Pittsburg, PA, pp. 191-195, Nov 1978.
22. M. I. Shamos, "Robust Picture Processing Operators and their Implementation as Circuits", Proc. Image Understanding Workshop, Pittsburgh, PA., pp. 127-129, Nov 1978.
23. P. M. Narendra, "A Separable Median Filter for Image Noise Smoothing," IEEE Trans. PAMI-3, pp. 20-29, Jan 1981.
24. A. Reeves and A. Rostampour, "Computational Cost of Image Registration with Parallel Binary Array Processor" (Correp.), IEEE Trans. PAMI-4, pp. 449-455, July 1982.
25. D. J. Delman, "Digital Pipelined Hardware Median Filter Design for Real-Time Image Processing", Proc. SPIE Conf. Vol 298, Real-Time Signal Processing IV, pp. 184-188, Sept 1981, San Diego, CA.
26. G. J. Yang and T. S. Huang, "Median Filters and Their Applications to Image Processing", Tech. Rpt. AD-8054507, School of Electrical Engineering, Purdue Univ., Dec. 1980.
27. D. C. Lai, J. Potenza, and K. Verfaillie, "Evaluation of Image Filters for Machine Classification," Optical Engineering (to be published).
28. H. C. Andrews and B. R. Hunt, Digital Image Restoration, Englewood Cliffs, N.J.: Prentice-Hall, 1977, pp. 126-132.
29. A. F. Lehar and R. J. Stevens, "Image Processing System for Enhancement and Deblurring of Photographs", Optical Engineering, Vol 23 (3), pp. 303-308, May/June 1984.

30. A. C. Bovik, T. S. Huang, and D. C. Munson, Jr., "A Generalization of Median Filtering Using Linear Combinations of Order Statistics", IEEE Trans. ASSP-31, pp. 1342-1350, Dec 1983.
31. J. B. Bednar and T. L. Watt, "Alpha-Trimmed Means and Their Relationship to Median Filters", IEEE Trans. ASSP-32, pp. 145-153, Feb 1984.
32. P. F. Velleman, "Robust Non-linear Data Smoothers: Definitions and Recommendations", Proc. Nat'l Acad. Sci., Vol 74, #2, pp. 434-436, 1977.
33. M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase", IEEE ASSP-24, pp. 58-63, Feb 1976.
34. P. E. Hart, "The Condensed Nearest Neighbor Rule", IEEE Trans. IT-14, No. 3 pp. 515-516, May 1968.
35. N. Weisstein, "Metacontrast", Visual Psychophysics (D. Jameson and L. M. Hurvich, Eds), New York: Springer-Verlag, 1972, Chap. 10.
36. G. L. Anderson and A. W. Netravali, "Image Restoration Based on a Subjective Criterion", IEEE Trans. SMC-6, pp. 845-853, Dec 1976.
37. S. A. Rajala and R. J. P. DeFigueiredo, "Adaptive Nonlinear Image Restoration by a Modified Kalman Filtering Approach" IEEE Trans. ASSP-29, pp. 1033-1042, Oct 1981.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PSYCHOLOGICAL CORRELATES OF PHYSIOLOGICAL INDICATORS OF STRESS RELATED  
DISORDERS: A SEARCH FOR STRUCTURE AND RELATEDNESS

Prepared by:	Dr. Charles L. Lardent, Jr.
Academic Rank:	Associate Professor
Department and University:	Department of Management Auburn University at Montgomery
Research Location:	Leadership and Management Development Center Air University
USAF Research:	Major (P) Lawrence O. Short Major Mickey R. Dansby
Date:	10 August 1984
Contract No:	F49620-82-C-0035

PSYCHOLOGICAL CORRELATES OF PHYSIOLOGICAL INDICATORS OF STRESS-RELATED  
DISORDERS: A SEARCH FOR STRUCTURE AND RELATEDNESS

by

Charles L. Lardent, Jr.

ABSTRACT

The detrimental consequences of stress have emerged as a major concern among diverse professional domains: medicine, psychology, sociology, human engineering, organizational science, management, etc. Stress influences not only individual health, but is intermeshed with a host of interrelated variables that strongly affect individual and organizational performance. This study is an inaugural and exploratory investigation of the relatedness of numerous personality, motivational, and biographical measures to established physiological indicators of stress-related disorders among a sample of high-performing, senior-level military officers and their spouses. The findings, only a select few of which can be presented here due to space limitations, reveal an unmistakable and significant complex of relationships as regards physiological and psychological indicators of stress. Several very significant findings emerged that have special relevance to the military. First, the conventional views and implications of the Type A Behavior Pattern are not entirely appropriate nor applicable to senior-level military officers. Second, an officer's rate of promotion is related to the ratio of total serum cholesterol divided by high-density lipoproteins, but the relationship is not linear. Third, there is a clear, seemingly inherent, difference between males and females on both psychological and physiological correlates of stress.

### ACKNOWLEDGEMENT

First of all, the author would like to thank the USAF Systems Command, the USAF Office of Scientific Research, the Air War College, the Leadership and Development Center, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to conduct a research project that is not only interesting and stimulating but one that also is significantly relevant and has far-reaching implications for the health and well being of senior-level military officers and their spouses.

Second, many people have influenced the successful culmination of this rather massive study: The entire Air War College (AWC) class of 1984, the College faculty, and the spouses of these officers. Specifically, a number of individuals deserve to be identified by name and thanked for their advice, support, and interest. They include Dr. (Col) Raymond G. Troxler of the USAF School of Aerospace Medicine; Dr. (Col) David O'Mara, Hospital Commander, and Major Roger Margotto, Chief of the Clinical Laboratory, both at Maxwell AFB; Chaplain (Lt Col) George Updegrove and Lt Col Robert Fowler of the AWC Department of Leadership and Management; Mr. Jesse Barron, Chief of the Statistical and Modeling Branch of the Air University Computer Center; and Majors Mickey R. Dansby and Lawrence O. Short of the Air University, Leadership and Management Development Center. Only through their efforts and support was it possible to bring this study to fruition. Finally, I want to express my appreciation to Mr. Glenn Donald for his expertise in Word Processing.

## I. INTRODUCTION:

Within the last ten to fifteen years, the role of psychological dimensions of health and illness has emerged as an area of great interest and concern. Among both researchers and practitioners, it has assumed a level of respectability never before attained. However, there is nothing novel about the notion of psychological involvement in physical well-being and illness. As long ago as 1936, Menninger and Menninger (21) discerned a relationship between psychoanalytic observations and cardiac disorders. Dunbar (8), Arlow (2), Kemple (16), and Alexander (1), all functioning within the psychoanalytic tradition, reported the relationship between "mind and body." Kemple (16), e.g., found that Rorschach protocols of patients with Coronary Heart Disease (CHD) revealed a strong and consistent pattern of aggressiveness and a need to control and dominate.

Today, it is generally recognized that psychological factors play important roles in most, if not all physical disorders, not just ulcers, insomnia, colitis, headaches, CHD, and a few other specific disorders. In fact, contemporary estimates are that between 70 - 90% of all so-called physical disorders have emotional or stress-related origins (26,12). "Stress" is a generic, nebulous, and loosely used term. In spite of extensive research, and notwithstanding a popular usage of the term which implies a knowledge of what we are dealing with, no one knows exactly what stress is; neither is there any common agreement on how to define stress. Selye (24,25), e.g., defines stress as the "nonspecific (i.e., common) result of any demand (stressor) upon the body, be the effect mental or somatic." This means that the biochemical response to any and all stressors (pleasant or painful) is qualitatively the same. Thus, Selye focuses on the adaptational capacity of the organism, and the nature of the stressors is largely irrelevant. In contrast, Lazarus (18) holds that an organism's cognitive appraisal of the nature of the stressors is an important variable. In fact, Lazarus (18) found it infeasible to try to operationally define stress. Instead, he holds that stress is a generic term that includes stressors (i.e., the stress stimuli), the ultimate outcomes or consequences, and the entire range of

intervening processes. Thus, in Lazarus' view, stress is not a stimulus, a response, or an intervening variable but, rather, an overall collective term that includes psychological, physiological, sociological, genetic, and other related phenomena. Lazarus' view is the one espoused by this author and is the one on which this study is predicated.

---

Insert Figure 1 here

Figure 1 portrays the author's composite schematic model of what loosely can be described as the stress phenomenon. The model presents the interdisciplinary nature of stress. Further, it illustrates why stress is not purely a medical problem, or a psychological one, or a physiological one, or one that stems from any single source. Rather, stress is multifaceted and extremely complex. Can the model identify the point at which stress arises? This is a moot and unanswerable question from an objective stand-point. From a subjective standpoint, the phenomenological experience of stress itself exists somewhere between external limits of the stressor domain and those of the ultimate consequence domain (3). Breznitz and Goldberger (3) suggest that, while the individual's experience of stress is the most germane factor in dealing with stress conditions, the stress experience itself exists outside the sphere of objective inquiry.

## II. OBJECTIVES OF STUDY:

Research into the causes, nature, treatment, and consequences of stress is accelerating at a rapid rate. Still, there remains vast unexplored areas and, even where research has taken place, incomplete and inconsistent findings have emerged. Nonetheless, even in terms of what is known today, the panorama of stress is so huge, complex, curvilinear, dynamic, and multi-disciplinary that it is infeasible (if not impossible) to construct an operational model that integrates all the domains depicted in Figure 1. Such a model can be abstractly conceptualized as an ideal, but it cannot yet be operationalized.

This study undertakes the primary task of exploring the relationships between aspects of the "individual psychological differences" block of the mediating domain and the "physiological" block of the

intermediate outcome domain. Additionally, the study also addresses, at least peripherally, the Type A Behavior Pattern (TABP) of the "behavioral" block of the intermediate outcome domain and several biodata factors within the "individual genetic and behavioral differences" block of the mediating domain. The findings generated by this study are voluminous, and, for that reason, only a small fraction of the results can be reported in this "final report of research effort." Hopefully, it will be possible to pursue this research effort, for the implications arising from the analyses of the two populations of interest (i.e., the Air War College students and faculty and their spouses) are not only interesting, significant, and useful, in several ways they call into question the conventional ideas about the generalizability of the role of the TABP, serum cholesterol (CHOL), high-density lipoproteins (HDL), and other indicators of stress prone outcomes.

### III. CRITERION MEASURES:

The overall study includes fourteen criterion variables for the AWC students and faculty, but only twelve for their female spouses. Because of measurement procedure problems, systolic blood pressure (SBP) and diastolic blood pressure (DBP) readings for the females were not sufficient to be included in the analyses.

---

#### Insert Table 1, Part D

---

Recent research has established that total serum cholesterol (CHOL) is not the independent and univariate precursor of coronary heart disease (CHD) it was once thought to be. Instead, it now appears that there are moderating components or fractions of CHOL that actually mitigate against CHD, while other components or fractions promote CHD. Specifically, the relationship between CHOL and its fractional component of HDL (i.e., CHOL/HDL) is the single best non-invasive predictor of actual CHD as documented by coronary angiography and other diagnostic medical techniques (14, 20, 28, 29). The greater the ratio of CHOL to HDL, the greater the risk of the existence of atherosclerosis and ultimate CHD. Conversely, low ratios of CHOL to HDL are associated with decreased risks and an increase of good health in general.

While all of the criterion variables listed in Table 1 were used in the study, because of space limitations, the CHOL/HDL ratio will be reported as the principal criterion in this "final report of effort." This ratio was selected for inclusion as the major criterion not because of its high level of statistical significance or its high degree of relatedness. Other criteria actually revealed greater significance and relatedness. Rather, the decision was based on the importance of the ratio in current research efforts, and on the fact that other Senior Service Colleges are making use of the ratio in their assessment programs. Thus, a basis for comparison will be possible between studies at the Army War College, the National Defense University, and the Air War College.

#### IV. PREDICTOR VARIABLES:

1. Before specifically addressing the predictor variables used in this study, it is essential to look briefly at some selected relevant research. Only by doing so can the foundation and rationale for the present research be put into perspective.

Jenkins, Hames, Zyzanski, Rosenman, and Friedman (14) conducted a study in which total serum cholesterol was regressed on the scales of the California Personality Inventory (CPI) (11). The sample consisted of 34 California firemen. Jenkins et al (14) were astounded to find relationships between cholesterol and psychological constructs. Four of the CPI scales were significantly related to serum cholesterol. Upon regression analysis, two scales - Socialization (this is not sociability) and Self-acceptance - yielded a multiple R of .66 and a corresponding  $R^2$  of .44. However, a replication of the study using 152 male supermarket managers and workers did not yield results as significant as those of the earlier study. While Socialization did emerge as a significant predictor variable in the second study, age, rather than psychological constructs, was the most significant. Jenkins et al (14) tentatively ascribed these differences in results to heterogeneity within the second sample and to restriction of range within the first. An extensive search of the literature reveals that the Jenkins et al (14) study is the only published research relating psychological constructs to serum lipids.

Friedman and Rosenman (10,23) originally described the Type A Behavior Pattern (TABP) and concluded that it is composed primarily of excessive drive, heightened sense of time urgency, and competitiveness. Later, Jenkins (13) concluded that the available evidence confirms a positive relatedness between TABP and CHD indicators, including medically confirmed atherosclerosis. However, at least two studies have shown no relationship between TABP and CHD indicators (6,7). The TABP is usually assessed in one of two ways: the Structured Interview (22) or the Jenkins Activity Survey (JAS) (15).

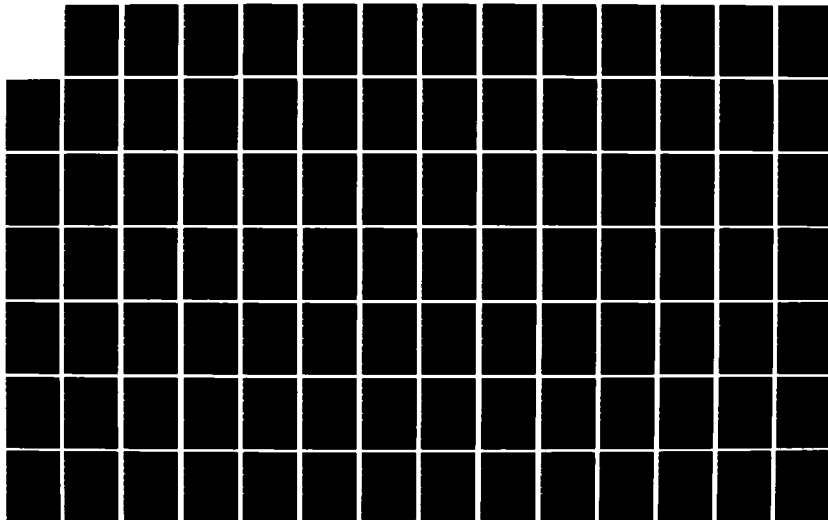
Jenkins, Zyzanski, and Rosenman (15) have concluded that the TABP is relatively independent of traditional CHD risk factors such as blood pressure, serum cholesterol, cigarette smoking, etc. They also conclude that the TABP is independent of existing psychological constructs within psychological theory. However, Krug and Johns (17) have shown that the Sixteen Personality Factor (16PF) questionnaire (4,17) accounts for 35% of the variance in the four scales of the JAS, with canonical correlation coefficients ranging from .47 to .75. Then, in view of the fact that there is a 20-25% error variance in the JAS, and that the 16PF is limited by its own reliabilities, Krug and Johns (17) conclude that personality may be the single largest factor influencing coronary prone behavior. Thus, it appears that the TABP is indeed related to (if not predicated on) psychological factors. After all, behaviors do not emerge out of nothingness. The point is, while inconsistencies do exist, there is sufficient evidence to suggest relationships between personality variables, the TABP, biochemical indicators of stress, and overt stress reactions themselves.

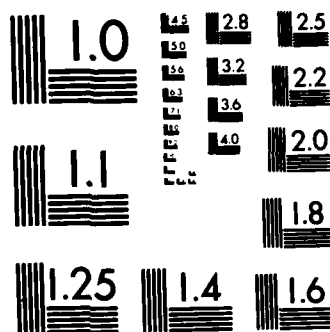
However, there is a chink in the TABP paradigmatic armor: not all (or even most) TABP individuals develop overt CHD! Based on this fact, and on the assumption that the TABP is obviously not the independent and univariate (or invariate) measure Jenkins, Zyzanski, and Rosenman (15) originally thought it to be, McCranie, Simpson, and Stevens (19) conducted a study in which the TABP was moderated by "field independence" versus "field dependence" (31). Individuals characterized by the TABP, but who are field independent, had far lower levels of CHOL, TRIGL, CHOL/HDL ratio, and LDL than those who also have the TABP but who are



AD-A154 337 UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM / 0/13  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

UNCLASSIFIED W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480 F/G 5/1 NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

field dependent. Thus, it appears that a person's cognitive style and phenomenological frame of reference (i.e., field independence versus field dependence) is at least as important as, and interacts with, the TABP in promoting or mitigating health risks.

When personality factors and ability factors are held constant, there remains a wide range of variability in human behavior. This is the domain of dynamic psychology, and a substantial part of this domain is explained by motivation. Personality (temperament) remains relatively stable over time; motivation, however, is more volatile and far more difficult to assess. The Motivational Analysis Test (MAT) (5,27) is the only objectively administered psychometric device available today that assesses the broad range of motivational structures. A survey of the literature reveals no evidence that motivation in general or the MAT in particular has been used in stress research.

In the present study, both the 16PF and the MAT are used as the major predictor devices. The TABP measures were computed using regression weights applied to the 16PF primary factor scores (17); they were not derived from the JAS or from the Structured Interview technique. Thus, for the TABP there is an obvious sacrifice in discriminatory power in the interest of feasibility and efficiency. In addition, for the AWC students and faculty, eight demographic or biographic factor categories were used; these generated 25 non-metric variables. For their spouses, however, only three categories were used (generating 13 non-metric variables), since five of the categories dealt with military unique activities or events.

---

Insert Table 1, Parts A, B, & C

Space does not permit a discussion of the predictor devices beyond identifying the scales contained in the devices. The 16PF consists of 16 primary personality factor traits which can be thought of as the basic building blocks or the source traits of personality. There are two validity scales: faking good and faking bad. The 16 primary factors combine in various ways to produce "second-order" factors or surface traits of personality. Seven of these were used in the study.

Over the years, the 16PF has been validated on various criterion groups, e.g.: leaders, teachers, accident victims, etc., (4,17). Of the almost unlimited number of criterion group norms, 23 were incorporated in parts of this study, but most yielded insignificant or only moderately significant results.

The MAT consists of 10 primary motivational structures, each of which yields four separate scores reflecting four separate views of each particular motive structure: unintegrated, integrated, total, and conflict. The unintegrated score is the primitive, mostly unconscious strength of the motive. The integrated score is the mostly conscious, social expression or actualization of the motive. The total score reflects a theoretical attempt to measure the combined motivational effect of the interaction of the unintegrated and integrated components. The conflict score is a measure of the dissidence or incongruity between unintegrated and integrated components. In addition, the MAT yields five composite measures, but they have not been adequately studied, and they did not emerge in this study as important predictor variables. Also, recent studies using the MAT have produced eight second-order motivation factors, and they are included in this study.

For both the 16PF and the MAT, the normative mean and standard deviation on each variable, for the North American population, are 5.5 and 2 respectively. This method of standardized scoring is called STEN (standard ten) scoring and yields a range of one to ten. The validities and reliabilities are reported in the respective 16PF and MAT handbooks (4,5).

#### V. THE POPULATIONS:

The study addresses two distinct populations: (1) the AWC student and faculty members and (2) the spouses of the AWC students and faculty. Over 93% of all eligible respondents participated. Three female AWC students and two male spouses were eliminated from the study since their numbers did not comprise sufficient size groups for analyses. Generally, the students and faculty totaled about  $N = 220$ , and the spouses about  $N = 163$ . However, these are only approximate figures since all

subjects did not participate in all assessment events, and, as mentioned, difficulties emerged in obtaining complete blood pressure readings.

The AWC student and faculty population comprises a unique and very selective group, and the generalizability of the findings presented in this study would be of questionable validity, except as applied to other senior service college populations and possibly to other groups of similarly high-performing, senior-level military officers. The average age of the male respondents was 43, with a range of 38 to 52 and a standard deviation of about 2. The group is extremely intelligent, with over 90% having at least one graduate degree. As a group, the respondents are high-performers, comprising about the top 15% of all officers within their group of contemporaries. Since all officers must meet physical and medical standards, the group as a whole would naturally be far "healthier" than the average American male population within the same age range. Intuitively, however, one would conclude that the AWC student and faculty population could be characterized by the TABP.

The female spouse population, while certainly not representative of the American population at large, is probably closer to the normal female population than their husbands are to the male. As a group, they also tend to be very bright but are not as highly formally educated as their husbands.

#### METHODOLOGY:

Over a period of about one month, beginning in January 1984, blood pressure was taken and blood serum collected from each AWC student, faculty member, and spouse who chose to participate. Over 25 blood chemistry measures were generated using automated analyzers; however, only nine were actually used in this study, and three additional ones were calculated by the author. These measures, which have been previously identified, comprised the criterion data domain. During the same period of time, both the 16PF and the MAT were administered to the two populations.

Since this study can best be described as an exploratory search for structure and relatedness, no formal hypotheses are presented. The domain of this study is virtually unexplored, with the only even

remotely related previous research being that of Jenkins et al (14) using the CPI. A wide range of statistical techniques were used in analyzing the data, including: correlational analysis, tests for differences between means (t-tests), regression analysis, analysis of variance, and canonical correlation analysis. Because of space limitations, in the next section only a very limited sample of results of analyses will be presented.

#### SELECTED RESULTS OF ANALYSES:

##### VI. DIFFERENCES BETWEEN MEANS:

In a search for pattern differences, a sizable number of t-tests were performed using both real and artificial dichotomies. Because of space limitations, only three will be reported here. Table 2 portrays the differences in means on both predictor and criterion variables that are significant at the  $p < .10$  level for rated officers versus non-rated officers.

#### Insert Table 2

No great surprises or dramatic findings emerged as regards the physiological criterion variables. Only one personality variable and one motivation variable were significant at the  $p < .01$  level: Tough Poise (QIII) and Mating-Unintegrated (Ma-U). The only significant criterion variable is diastolic blood pressure (DBP), with the rated officers being significantly but not dramatically lower. Nonetheless, an examination of Table 2 reflects an unmistakable pattern of personality and motivational differences, between rated and non-rated officers, that is statistically significant if not quantitatively substantial.

#### Insert Table 3

Table 3 provides far more intriguing results with very interesting and revealing implications. It reflects a 3-way comparison of officers having only primary zone (PZ) promotions, those with one secondary zone

(SZ) promotion, and those with two or more SZ promotions. In retrospect, it would have been useful to have determined the timing and the grade level at which the SZ promotions occurred; however, the author had no a priori basis for inferring the importance that later emerged. What is so remarkable are the differences between Group A and Group B, Group B and Group C, but not between Group A and Group C. Given that significant findings should arise, one would naturally expect linearity along the psychological and physiological variables as the number of SZ promotions increases from none to a maximum of three, but this is not the case. For several of the significant variables the pattern is curvilinear.

It is interesting that the officer group with more than one SZ promotion (Group C) is significantly more suspicious (L+) than either of the other two groups. Elevated L (although 6.4 is not excessively high) can be best understood as "anxious insecurity," and it is interesting to speculate why the mean for these "super-fast burners" is almost one-half standard deviation above the means of the other two groups, as well as the general population.

The career motive also presents an interesting pattern. Those officers having no SZ promotions (Group A) apparently have the highest level of career motive achievement or actualization (Ca-I). At first glance this seems to be a non sequitur; one would expect Group C to be the highest on Ca-I. However, when career conflict (Ca-Conf) is added to the picture, the pattern begins to make perfectly good sense. Career motive conflict (Ca-Conf) is a function of the difference between Ca-U and Ca-I. Those officers having two or more SZ promotions (Group C) show higher Ca-Conf than Group A. Similarly, Group B is almost at the same level as Group C. What this pattern suggests is that Group A, compared to Groups B and C, is achieving career motive aspirations; their aspirations are simply less ambitious. There is less tension or conflict associated with further career achievement motivation. In a sense, then, career drive (Ca-U) is not as strong a motivational force for Group A as it is for Groups B and C, mainly because Group A is more career satisfied (or career resigned) than Groups B or C. Implicit in

this pattern is a picture of the relationship between career and potential stress that is far more complex than the traditional univariate model of job stress giving rise to a direct stress response.

The assertiveness structure also deserves comment, for it also presents a somewhat anomalous pattern along the conflict dimension. Group C shows the greatest As-Conf; Group B, surprisingly, is the lowest; and Group A is in the middle, and is not significantly different from Groups B and C. Again, the pattern is a curvilinear one and not what would be rationally expected.

The biochemical criterion variables are similarly interesting and confirm the psychological patterns. For example, Group C has the highest CHOL/HDL ratio, and, more important, it is within an elevated CHD risk range (20). Based on rational speculation, then, Group A should be the lowest, and Group B should be somewhere in between; but this is not the case. Group B is definitely the lowest on COL/HDL ratio. The pattern that emerges is a curvilinear one. It suggests that Group B, those officers with one SZ promotion, is relaxed, rather "laid-back," and comparatively self-satisfied and unstressed. Both Groups A and C tend to be under greater psychological and physiological stress, with Group C being the higher of the two, although the difference is not statistically significant except for Suspiciousness (L+) and Ca-Conf.

---

Insert Tables 4 and 5

---

Table 4 reflects the significant differences in means for the males when the male research population is artificially dichotomized on the CHOL/HDL ratio criterion variable. Table 5 does the same for the females. In order to highlight the structural differences, the respondents within one-half standard deviation of the respective research population means were eliminated from the analysis. Thus, Groups A are at least one-half standard deviation below the respective research population means, and Groups B are at least one-half standard deviation above the two respective means on the CHOL/HDL ratio.

Tables 4 and 5 reveal a number of interesting findings; however, space limitations preclude a full discussion of them. It is noteworthy,



though, to see the obvious interrelatedness of almost all of the physiological variables. Space does not permit the inclusion of intercorrelation matrices which would better highlight those relationships. Table 4 dramatically, though admittedly artificially, reveals the importance of CHOL/HDL ratio as a correlate of a sizable number of personality and motivational variables, as well as of other physiological criterion variables; but this general pattern applies primarily to males. Table 5 reveals that, for females, personality is apparently unimportant and unrelated to the CHOL/HDL ratio. Interestingly, however (though unreported in the final report), female personality scores are definitely related to HDL but not always in the expected direction. Also, in contrast to the male pattern, the affection motive (SW) is extremely important for females (but not for males) as it relates to the CHOL/HDL ratio.

#### VII. REGRESSION ANALYSIS:

Can the criterion variables be predicted from the psychological and biographical data? With Multiple Rs ranging from .23 (Glu) to .48 (UA) for males and .27 (CHOL/HDL ratio) to .60 (C1) for females, the answer is: To a remarkably surprising extent! For reasons cited earlier, only the CHOL/HDL ratio will be presented in this report.

---

#### Insert Table 6

Table 6 portrays the results, for males, of regressing the CHOL/HDL ratio on a number of psychological and biographical variables that were originally selected based on an inspection of intercorrelations between variables. Stepwise regression techniques were used, and the variables shown in Table 6 comprise the best predictor equation. The table invites a great deal of interpretation and discussion that space does not permit; however, what is especially noteworthy is the degree to which "more than one SZ promotion" contributes to the CHOL/HDL ratio. This finding operationalizes the promotion rate differences reflected in Table 3.

---

#### Insert Table 7

Table 7 depicts for females what Table 6 does for males. What emerges from a comparison of the two regression equations is a dramatic insight into the different patterns exhibited by males and females. The differences are remarkable. They would appear even more profound were it possible to present other regression analyses. Obviously, the CHOL/HDL ratio does not have the same implications for females as it does for males. For example, as females tend to express assertiveness (As-I), there is a corresponding increase in the CHOL/HDL ratio. However, as males actualize underlying assertive drives (As-Tot), there is a corresponding decrease in the CHOL/HDL ratio. Also, as females become dominant (E) in their personality orientation, there is a corresponding increase in their CHOL/HDL ratio. While dominance (E) did not appear in the regression equation for men, there is nonetheless an  $r = -.17$  correlation between that predictor variable and the CHOL/HDL ratio. That is exactly the opposite of the females. The implications of these findings are intriguing and portentous. An obvious research question is, are these findings unique to senior-level military officers, or are they generalizable to other groups such as civilian middle-level managers?

#### VIII. ANALYSIS OF VARIANCE:

##### Insert Table 8

Table 8 depicts the results of a one-way analysis of variance, for males, when the TABP (as estimated by the 16PF) is moderated by the 16PF second-order factor of Independence ( $Q_{IV}$ ). The cut-off points for both the measures were inserted at the medians, thus producing four cells: TABP-Independent, TABP-Dependent, TBBP-Independent, and TBBP-Dependent. Of course, in no way should it be inferred that those officers characterized as having the TBBP are actually Type Bs. They are characterized as such merely as a result of an artificial dichotomy at the median points of the distribution. The criterion variable is the CHOL/HDL ratio. This procedure was suggested by the research design

used in a study previously cited (19); however, the predictor instruments are different, and it should be tenuous to equate Witkins "field independence" (30) with the 16PF measure of Independence (Q<sub>IV</sub>). Also, in the earlier study, multiple criteria were used; here the only criterion variable is the CHOL/HDL ratio.

In spite of these methodological differences, however, the pattern shown in Table 8 is essentially the same as the one found in the earlier study (19), thus, confirming the importance and the obvious influence of cognitive and operational frames of reference as moderators of the TABP. Table 1 reveals that the AWC students and faculty are very independent, with a mean score of 6.8. Also, they clearly reflect the TABP, with a mean score of 7.1. The correlation between the TABP and the CHOL/HDL ratio is  $r = -.14$ , and with HDL it is  $r = .21$ . Under the conventional view of the TABP, these relationships do not make sense; they do not agree. However, under the paradigm of the TABP moderated by independence/dependence, they do make sense, and the implications are far reaching. The TABP for senior-level military officers apparently does not operate quite in the same way that it does for the population at large, and, at least in part, this difference is explained by "field independence." For females, however, again the results are different, are not significant, and, therefore, are not reported due to space constraints.

#### IX. Conclusions and Recommendations:

At the highest level of abstraction, what emerges from this study is a clear picture of the complexity of "mind-body" interrelatedness. In many instances the relatedness and structures which emerge are quite dramatic and unexpected. Certainly, this study has opened vistas that should lead to a fertile field of research in promoting a better understanding of stress phenomena in the military. Some of the findings present real enigmas. For example, while not included in this final report of effort due to space limitations, there emerged, for females, a Multiple R of .60 between the psychological predictor domain and Chloride (Cl). Even when the Multiple R is "shrunk" to approximate what cross-validation would yield, it only drops to  $R = .54$ . The question is: WHY?

Chloride is an electrolyte, the major extracellular negative ion, and it is involved in maintaining water distribution, osmotic pressure, and ion balance with Sodium and Potassium. Apparently, however, especially for females, it is also very significantly related to psychological constructs. Consultations with medical and laboratory professionals produced only amazement and no answers (28,29). The question of Chloride is cited here merely as an example of several enigmas which emerged and highlights the need for more intensive and extensive research.

At a somewhat lower level of abstraction, this study illustrates the importance of personality and motivational constructs in context with each other, along with relevant biographic factors. Personality structures do not exist in isolation, and the additional explanatory power provided by the MAT documents the interaction between personality and motivation.

This study identifies and documents important distinctions between males and females. The two research populations respond quite differently, both psychologically and physiologically, on a sizable number of variables. When several of the findings were shared with Dr. George Troxler (28), he expressed intense interest, astonishment, and satisfaction. According to Troxler, several studies have identified

differential cholesterol levels between males and females, but an exhaustive search of the literature revealed no evidence of attempts to explore interrelational behaviors and other biochemical pattern differences. Yet, Troxler reports his own research has uncovered some intriguing biochemical anomalies between the sexes in their responses to stress stimuli. This current study, according to Troxler, is the only known attempt to explore psychological influences of differential biochemical responses. One or two of these differential responses should be highlighted.

The study provides evidence of the importance of "Mating" as measured by the MAT. Mating-U is the sex-drive itself; Ma-I is the actualization or satisfaction of the sex-drive motive; Ma-Conf is the measure of the dissidence between Ma-U and Ma-I. For males, as Ma-U and Ma-Conf increase, a healthy drop in the CHOL/HDL Ratio results; both are statistically significant at  $P < .01$ . However, while not significant at conventional levels ( $P = .30$ ), as Ma-I increases there is a corresponding increase in the CHOL/HDL Ratio. For females, however, as Ma-Conf increases, there is a corresponding increase in the CHOL/HDL Ratio ( $P = .07$ ). A similar pattern exists for Ma-U, but it is not significant at conventional levels ( $P = .25$ ). Also, as Ma-I increases there is a drop in the CHOL/HDL Ratio, but it is not significant at conventional levels ( $P = .21$ ). Thus, what emerges is a picture of differential response wherein the sex-drive itself, for males, lowers the CHOL/HDL Ratio, while sex-satisfaction is associated with an increase in the Ratio. It seems to be exactly the opposite for females, with sex-satisfaction being associated with a decrease in the CHOL/HDL Ratio.

The Assertiveness motive also presents an interesting sex-difference anomaly. For both males and females, as As-U increases there is a decline in the CHOL/HDL Ratio. As males express their assertiveness (As-I), there is also a decline in the ratio; however, for females, as As-I increases, there is a corresponding increase in the CHOL/HDL Ratio. Thus, it appears that males and females handle assertiveness differently, and the consequences are less favorable for the females. The same may be said for the personality factor: Dominance (E). For males,

as E increases, there is a decline in the CHOL/HDL Ratio. For females, though, as E increases, the CHOL/HDL Ratio increases. This suggests that dominance in males, at least in the male population of interest, is a very desirable attribute. For females, however, it is a less favorable attribute as far as CHD risks are concerned.

There are so many research findings to emerge from this study that it is difficult to decide what to highlight within the allowable space. However, the Affection motive screams for comment since it produced the highest conflict scores (SW-Conf) and also the highest unintegrated scores (SW-U) for both males and females (see Table 1B). A recurrent theme within the military community is the existence of a high need to receive affection but an inability to give it. This study documents that theme, but it is not clear how the motive works in influencing the biomechemical variables. Table 5 suggests that Affection is extremely important with respect to the CHOL/HDL Ratio, but none of its dimensions emerged as significant predictors in the regression analysis. Table 4 suggests that, for males, Affection is not at all important with respect to the CHOL/HDL Ratio; yet, SW-Tot emerged as a minor but significant predictor as shown in Table 6. Obviously, much more research is needed to understand how this important motive works.

Perhaps the single most puzzling finding to emerge during the early stages of this study was the negative relationship between the TABP and the CHOL/HDL Ratio. Upon subsequent analysis, however, the supposed monolithic and univariate nature of the TABP was questioned when it was moderated by "field independence." Thus, at least for the senior-level military officers who comprised the male research population of this study, the TABP, as a discreet and independent construct, is a less viable indicator of the CHOL/HDL Ratio and CHD risks than it is for the general population.

The single greatest risk factor for CHD, as well as most other disorders, is age; and the group at highest risk are males in the age range addressed by this study. Unfortunately, nothing can be done about age, but a great deal can be done about other risk factors that slow down and mitigate the effects of age. This study, in addressing a

virtually unexplored domain, has opened vistas with respect to the psychological correlates and influences of physiological indicators of stress-related disorders. The findings, only a few of which are presented here, establish and/or confirm the complexity and interrelatedness of many factors contributing to the stress phenomenon.

The military officer in the 40-50 year old age range is particularly vulnerable to a number of health risks, especially CHD. Those senior level officers selected to attend senior service colleges are prime candidates for health disorders, yet they are the competent high performers identified as our future general officers and senior military leaders. Aside from the human costs, the investment placed in these officers is astronomical. It is incumbent on both the military services and on the senior-level officers themselves to gain a better understanding of the sources of psychological and physiological health risks and to learn to mitigate their destructive effects. While this present research effort has focused on the individualized consequences of stress, the "costs" of stress are not limited to the individual; there are also organizational "costs." Fielder (9), for example, has shown how insidiously dysfunctional individual and organizational stress can be in achieving organizational objectives in a large military organization. However, the overall stress phenomenon is so complex, interrelated, and pervasive of multiple disciplines that, at our present level of knowledge, no single research effort can capture its totality. Far more "limited domain" research needs to be undertaken, and the author would welcome the opportunity to continue his research efforts within the senior-level military officer population.

Of course, senior-level officers and their spouses do not spontaneously become candidates for health problems. The groundwork is laid years earlier. For this reason, it needs to be pointed out that similar studies need to be undertaken at other levels: Command and Staff, Squadron Officer, NCO, etc. Only in this way can baselines be established. Moreover, while a lot is said about the "whole person," very little is actually accomplished in promoting understanding and developing the "whole person." Hopefully, the AWC experience on which this study is based is a meaningful step in addressing this deficiency.

# BIBLIOGRAPHY

1. Alexander, F. (1950) Psychosomatic Medicine. New York: Norton
2. Arlow, J. A. (1945) Identification Mechanisms in Coronary Occlusions. Psychosomatic Medicine, 7, 195-209
3. Breznitz, S. and Goldberger, L. (1982). Stress Research at the crossroads. In L. Goldberger and S. Breznitz (Ed.), Handbook of Stress: Theoretical and Clinical Aspects (pp. 3-6). New York: The Free Press.
4. Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. (1970). Handbook for the 16PF. Champaign: IPAT.
5. Cattell, R. B., Horn, J. L., Sweeney, A. B., Radcliffe, J. A. (1964). The Motivational Analysis Test ("MAT"). Champaign: IPAT.
6. Dimsdale, J. E., Hackett, T. P., Hutter, A. M., Block, P. C., and Catanzano, P. M. (1978). Type A Personality and Extent of Coronary Atherosclerosis. American Journal of Cardiology, 42, 583-586.
7. Dimsdale, J. E., Hackett, T. P., Hutter, A. M., Block, P. C., and Catanzano, D. M. (1979). Type A Behavior and Angiographic Findings. Journal of Psychosomatic Research, 23, 273-276.
8. Dunbar, F. (1943). Psychosomatic Diagnosis. New York: Hoeber & Harper.
9. Fiedler, F. E. (1982) Are Leaders an Intelligent Form of Life? The Role of Cognitive Processes in Leadership Performance. (Technical Report 82-1) Seattle: University of Washington, Department of Psychology.
10. Friedman, M. and Rosenman, R. H. (1974). Type A Behavior and Your Heart. New York: Knoph.
11. Gough, H. G. (1957). Manual for the California Psychological Inventory. Palo Alto: Consulting Psychologists Press.
12. Ivancevich, J. M. and Matterson, M. T. (1980). Stress and Work: A Managerial Perspective. New York: Scott, Foresman.
13. Jenkins, C.D. (1976). Recent Evidence Supporting Psychologic and Social Risk Factors for Coronary Disease. New England Journal of Medicine, 294, 987-994, 1033-1038.
14. Jenkins, C. D., Hames, C. G., Zyzanski, S. J., Rosenman, R. H. and Friedman, M. (1969). Psychological Traits and Serum Lipids. Psychosomatic Medicine, 31(2), 115-128.



15. Jenkins C. D., Zyzanski, and Rosenman, R. H. (1979). Jenkins Activity Survey. New York: Psychological Corporation.
16. Kemple, C. (1945) Rorschach Method and Psychosomatic Diagnosis. Psychosomatic Medicine, 7, 85-89
17. Krug, S. E. and Johns, E. F. (1984). Personality Factors underlying Coronary Prone Behavior Patterns. Champaign: IPAT (unpublished manuscript).
18. Lazarus, R. S. (1966) Psychological Stress and the Coping Process. New York: McGraw-Hill
19. McCranie, E. W., Simpson, M. E., and Sterns, J. S. (1981). Type A Behavior, Field Dependence, and Serum Lipids. Psychosomatic Medicine, 43(2), 107-116.
20. Majeski, E. J. (1978). Cardiac Risk Evaluation by Analysis of Lipoprotein cholesterol. Condenser (Bio-Science Labs), 9(1), 1-6.
21. Menninger, K. A. and Menninger, W. C. (1936) Psychoanalytic Observations in Cardiac Disorders. American Heart Journal, 11 10-21.
22. Rosenman, R. H. (1978) The Interview Method of Assessment of Coronary-Prone Behavior Pattern. In T. M. Dembroski, S.M. Weiss, J. L. Shields, S. G. Haynes, and M. Feinleid (Eds), Coronary-Prone Behavior. New York: Springer.
23. Rosenman, R. H., Friedman, M., Strass, R., Worm, M., Kositchek, R., Hahn, W., and Wherthessen, N. (1964). A Predictive Study of Coronary Heart Disease: The Western collaborative Group Study. Journal of the American Medical Association, 189, 15-22.
24. Selye, H. (1976). The Stress of Life. (2d ed.) New York: McGraw-Hill.
25. Selye, H. (1974) Stress without Distress. Philadelphia: Lippincott.
26. Shapiro, A. K., (1978) Placebo Effects in Medical and Psychological Therapies. In S. L. Garfield and A. E. Bergen (Ed.), Handbook of Psychotherapy: and Behavioral Change: An Empirical Analysis (pp. 369-410). New York: Wiley.
27. Sweney, A. B. (1969). Individual Assessment with the Motivational Analysis Test. Champaign: IPAT.
28. Troxler, R. G. (1984). Former Chief of Clinical Pathology, School of Aerospace Medicine, Brooks AFB, TX. Personal communications regarding unpublished Stress Research. Dr. Troxler is now in Private Practice in San Antonio, TX.

29. Troxler, P. G., Tolan, G. D., and Jackson, W. G. (1983) Coronary Angiography Predictors. Brooks AFB, TX: USAF School of Aerospace Medicine (unpublished manuscript).
30. Witkin, H. A. and Goodenough, D. R. (1977) Field Dependence and Interpersonal Behavior. Psychological Bulletin, 84, 661-689
31. Witkin, H. A., Oltman, P. K. Raskin, E., Kare, S. A. (1971). Manual for the Embedded Figures Tests. Palo Alto: consulting Psychologists Press.

FIGURE 1

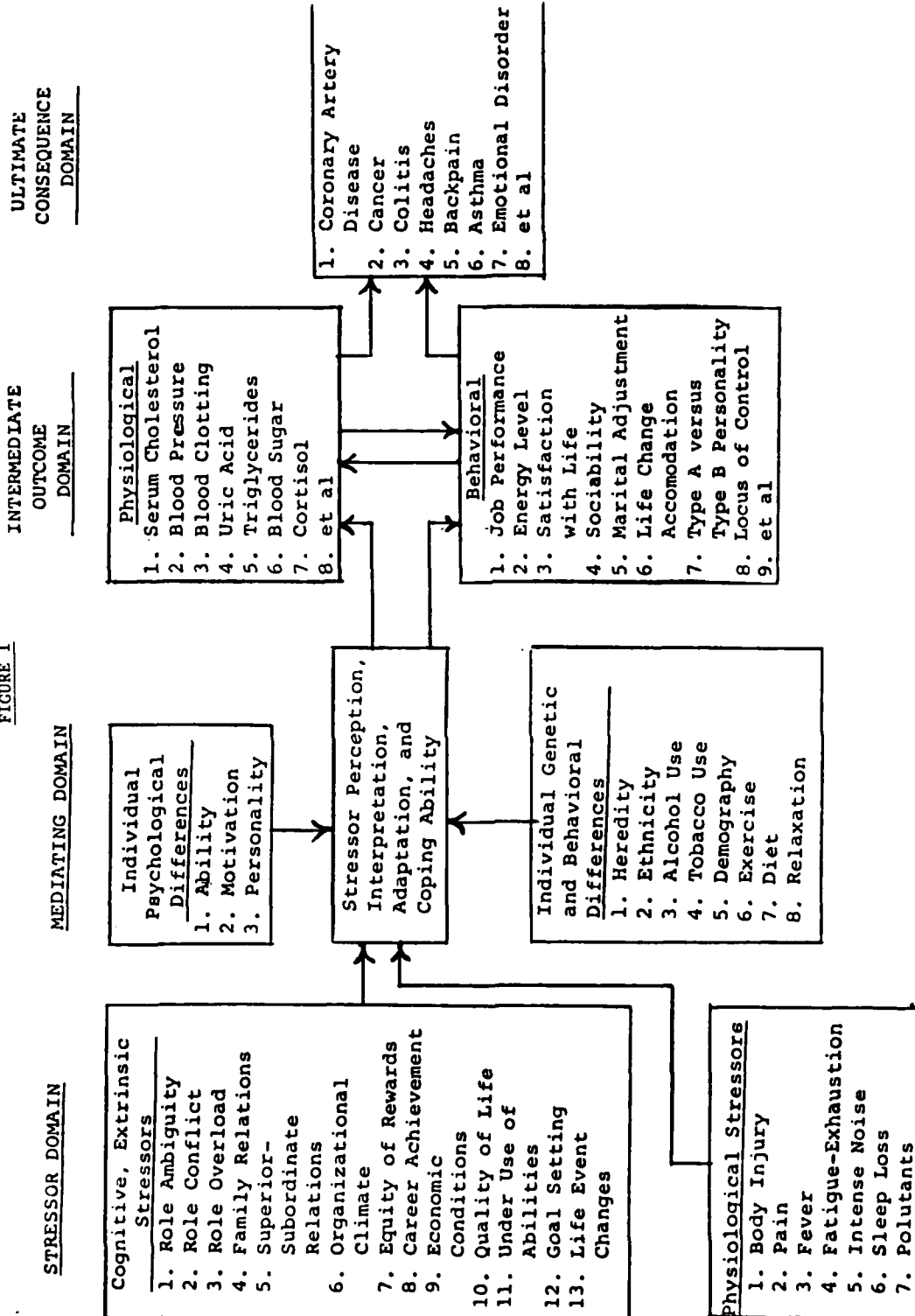


Table 1

Variable List with means and Standard Deviations. Except where noted, predictor variables are expressed in T-scores. For global American Population  $\bar{X}=5.5$  and  $SD=2$

I. Predictor Variables		Students & Faculty - Males		Spouses-Females	
A. Personality - 16 PF		N = 220		N = 162	
1. Primary source trait factors:		$\bar{X}$	SD	$\bar{X}$	SD
P1 (A)	Warm, outgoing, participating	5.3	2.1	5.8	2.2
P2 (B)	Intelligent	7.7	1.5	7.1	1.8
P3 (C)	Ego strength, calm emotionally stable	5.6	1.8	5.6	1.9
P4 (E)	Dominant, Competitive, Assertive	7.2	1.9	5.9	2.0
P5 (F)	Impulsive, happy-go-lucky	5.5	2.0	5.5	2.3
P6 (G)	Superego strength, group conformity, Persevering	6.4	1.8	6.0	1.9
P7 (H)	Bold, venturesome, active	5.9	2.2	6.0	2.1
P8 (I)	Emotionally Sensitive, dependent	5.0	2.0	5.6	1.7
P9 (L)	Suspicious, Jealous, dogmatic	5.7	2.0	5.6	2.0
P10 (M)	Imaginative, unconventional, bohemian	6.3	1.7	5.6	1.9
P11 (N)	Shrewd, Astute, Wordly	5.5	1.8	5.4	1.9
P12 (O)	Apprehensive, guilt prone	5.0	1.6	5.0	1.8
P13 (Q1)	Radical, experimenting, liberal	5.5	1.9	5.0	1.9
P14 (Q2)	Self-sufficient, resourceful	6.5	1.8	6.4	2.0
P15 (Q3)	Controlled, Ability to bind anxiety, precise	6.0	1.6	6.2	1.8
P16 (Q4)	Tense, fretful, free-floating anxiety	6.5	1.9	6.1	2.0
2. Validity (faking and distortion) measures:					
P17	Tendency to fake "bad"	5.3	1.0	5.3	1.1
P18	Tendency to fake "good"	6.4	1.6	5.7	1.7
3. Second-order, surface trait factors:					
P19 (Q1)	Extraversion	6.0	1.9	5.7	2.1
P20 (Q11)	Anxiety	5.8	1.6	5.4	1.7
P21 (Q111)	Cortical alertness (tough poise)	6.5	1.9	5.9	2.0
P22 (Q1V)	Independence	6.8	1.6	5.9	1.7
P23 (QV)	Discreteness	5.7	1.8	5.4	1.8
P24 (QVI)	Prodigal Subjectivity	5.8	1.9	6.0	2.0
P25 (QVII)	Superego, compulsivity, rigidity	6.3	1.7	6.0	1.7

Table 1 (cont.)

4. Criterion group measures:			
	$\bar{X}$	SD	SD
P26 Neuroticism	5.1	1.7	1.8
P27 Psychoticism	5.4	1.8	1.9
P28 Acting-out behavior	5.7	1.3	1.3
P29 Behavior control (similar to QVIII)	6.0	1.6	1.7
P30 (not used)	6.8	2.1	1.9
P31 Type A behavior Pattern (TAPP)	7.1	2.0	2.3
P32 Speed and Impatience	6.8	2.0	2.1
P33 Job Involvement	6.0	2.0	2.2
P34 Hard-driving and competitive	6.6	1.9	2.2
P35 Vulnerability to stress*	29.7	12.5	11.4
P36 Freedom from Accidents	5.6	1.1	1.3
P37 Teaching Effectiveness	6.7	1.8	1.9
P38 creative orientation	6.9	1.5	1.8
P39 Leadership-Elected	7.7	2.5	2.8
P40 Leadership-Effective	6.4	2.2	2.4
P41 Leadership-Technical	8.0	2.1	2.3
P42 Leadership role compatibility	6.3	1.4	1.5
P43 Attention to detail	5.2	2.4	2.5
P44 Attention to Rules and Regulations	4.9	1.8	1.9
P45 Academic Achievement Potential	7.6	1.6	1.9
P46 Predicted Auto Accidents-Married-10,000 miles*	.1	.4	.4
P47 Predicted Auto Accidents-Single-10,000 miles*	.5	.4	.4
P48 Marital Adjustment-husband only	6.9	2.2	-
P49 Marital Adjustment-wife only	-	-	2.4

\* Not Sten scores

Table 1 (cont)

MALES, N = 210										
B. Motivation - MAT										
Primary unintegrated(U), integrated(I), total(T), and conflict(C) factor scores; Global scores; and										
Second-order Factor scores for males and females										
1. Primary Factor Scores	$\bar{X}$				SD					
	U	I	T	C	U	I	T	C	I	C
M1,11,21,31(Ca) Career	6.0	5.4	5.7	6.6	1.7	1.7	2.3	2.2		
M2,12,22,32(Ho) Home-parents	4.9	5.3	4.5	5.6	1.8	1.4	2.1	2.2		
M3,13,23,33(Fr) Fear	4.2	6.7	5.4	3.3	1.8	1.9	2.6	2.3		
M4,14,24,34(Na) Narcism	6.5	8.0	8.4	3.9	1.7	1.5	1.7	2.5		
M5,15,25,35(Se) Superego	4.2	4.7	3.4	5.3	1.5	1.7	2.2	2.3		
M6,16,26,36(Ss) Self-Sentiment	4.9	6.1	5.1	4.8	1.7	1.3	2.2	2.1		
M7,17,27,37(Ma) Mating (sex)	5.5	6.6	6.1	4.6	1.7	1.6	2.2	2.2		
M8,18,28,38(Pg) Pugnacity, power	6.1	8.1	7.8	4.0	2.0	1.6	2.2	2.4		
M9,19,29,39(As) Assertiveness	5.2	5.2	4.5	5.9	1.8	1.8	2.4	2.2		
M10,20,30,40(Sw) Affection	7.1	5.0	6.3	7.7	1.6	1.8	2.4	1.9		
2. Global scores	$\bar{X}$				SD					
M41 Total Integration			7.8				1.4			
M42 Total personal Interest			7.2				1.1			
M43 Total conflict			5.5				1.3			
M44 Autism (optimism)			4.9				1.5			
M45 Information			9.0				1.2			
3. Second-order Factors										
M46 Self-Sentiment (ego)			5.5				1.4			
M47 Cooperativeness			3.7				1.4			
M48 Primitive comfort			6.5				1.3			
M49 Protestant work ethic			5.5				1.3			
M50 Sexual confirmation			5.5				1.3			
M51 Growth tension			5.1				1.3			
M52 Self-preeminence			5.4				1.3			
M53 Career tension			6.4				1.4			



Table 1 (cont)

C. Biographical variables: all non-metric values		Males N =	Females N =
V1.	Locale of birth or early years of life.		
D1	Criterion - other	29	17
D2	North	75	30
D3	South	64	52
D4	West	42	22
V2.	Size of locale (city) of birth		
D5	Criterion - other (more than 500,000)	33	12
D6	Less than 5,000	56	31
D7	Between 5,000 and 100,000	95	62
D8	Between 100,000 and 500,000	26	16
V3.	Ethnicity		
D9	Criterion (Black, Hispanic, Oriental, etc.)	6	2
D10	Northern Europe	169	84
D11	Southern Europe	13	9
D12	Slavic	7	13
D13	Europe - other	15	13
V4.	Military Rank		
D14	Criterion - Colonel (0-6)	105	-
D15	Lieutenant Colonel (0-5)	105	-
V5.	Rating Status		
D16	Criterion - None	108	-
D17	Pilot	87	-
D18	Navigator	15	-
V6.	Rate of Promotion		
D19	Criterion - No SZ promotion	88	-
D20	More than one SZ promotion	41	-
D21	One SZ promotion	81	-
V7.	Role orientation		
D22	Criterion - other than commander/LDR	110	-
D23	Commander/manager/leader.	100	-
V8.	Functional orientation		
D24	Criterion - other than operations	102	-
D25	Operations.	108	-



Table 1 (cont)

U. Criterion Variables	Males N = 204		Females N = 123	
	$\bar{X}$	SD	$\bar{X}$	SD
V 9 Systolic Blood Pressure (SBP)	123.0	11.5	-	-
V10 Diastolic Blood Pressure (DBP)	79.5	7.1	-	-
V11 Total Serum cholesterol (Chol)	210.8	45.0	196.8	50.2
V12 Triglycerides (Trigl)	115.6	79.1	87.3	46.3
V13 High-density lipoproteins (HDL)	42.3	12.8	49.4	13.6
V14 HDL Ratio (Chol/HDL)	5.3	2.0	4.3	2.2
V15 Glucose (Glu)	100.0	22.1	96.3	18.3
V16 Uric Acid (UA)	5.4	1.1	3.6	.9
V17 Sodium (NA)	142.7	3.8	141.5	3.5
V18 Potassium (K)	4.5	.6	4.5	.5
V19 Chloride (Cl)	104.9	3.3	104.8	3.2
V20 Alanine AminoTransferase (ALT)	33.3	13.7	27.2	21.7
V21 Low-density lipoproteins (LDL)	146.9	42.6	130.0	49.5
V22 LDL Ratio (LDL/HDL)	3.8	1.7	3.0	2.0

TABLE 2  
Significant Differences Between  
Rated and non-Rated officers on  
predictor and criterion variables.

	Non-Rated N = 108		Rated N = 94		t	$p \leq$
	$\bar{X}$	SD	$\bar{X}$	SD		
P1(A) Warmth	5.7	2.0	5.0	2.0	2.25	.05
P4(E) Dominance	7.0	1.8	7.5	2.0	-1.65	.10
P11(N) Shrewdness	5.6	1.8	5.2	1.8	1.90	.10
P21(QIII) Tough poise	6.1	2.0	6.9	1.8	-3.11	.01
P22(QIV) Independence	6.5	1.6	7.0	1.5	-2.26	.05
P23(QV) Discreteness	5.9	1.8	5.4	1.7	2.12	.05
P38 Creativity	6.7	1.6	7.1	1.4	-1.79	.10
M5(SE-U) Superego	4.4	1.4	3.9	1.6	2.14	.05
M7(Ma-U) Mating	5.3	1.5	5.9	1.8	-2.50	.01
M23(Fr-Tot) Fear	5.6	2.6	4.9	2.6	1.98	.05
M25(SE-Tot) Superego	3.6	2.2	3.1	2.1	1.64	.10
M27(Ma-Tot) Mating	5.8	2.2	6.4	2.2	-1.99	.05
M37(Ma-Conf)Mating	4.4	1.9	5.0	2.5	-1.99	.10
M44 Autism	5.1	1.5	4.7	1.6	2.01	.05
M50 Sex confirmation	5.4	1.3	5.7	1.3	-1.97	.05
V10 Diastolic BP	80.7	6.9	78.2	7.3	2.37	.05

2-tail test

All other differences were not significant at the  $p \leq .10$  level.

TABLE 3

Significant Differences Between Officers having no Secondary Zone promotions, one SZ promotion, and more than one SZ Promotion on predictor and criterion variables.

	A No SZ PROMO N = 88			B One SZ PROMO N = 81			C >One SZ PROMO N = 33			A/B		A/C		B/C	
	$\bar{X}$	SD		$\bar{X}$	SD		$\bar{X}$	SD		t	p	t	p	t	p
P9 (L)	5.6	2.0	Suspiciousness	5.5	1.9		6.4	2.0		--		-1.98	.05	-2.23	.05
P11(N)	5.7	2.0	Shrewdness	5.2	1.8		5.3	1.1		1.75	.10	--	--	--	--
P23(QV)	5.9	2.0	Discreteness	5.4	1.8		5.6	1.1		1.72	.10	--	--	--	--
P40	6.5	2.2	LDRSHP-Elected	6.0	2.3		6.8	2.3		--		--		-1.67	.10
M11(Ca-1)	5.7	1.7	Career	5.1	1.8		5.3	1.7		2.07	.05	--	--	--	--
M13(Fr-1)	6.9	1.9	Fear	6.4	1.9		6.8	1.9		1.95	.05	--	--	--	--
M23(Fr-TOT)	5.5	2.5	Fear	4.9	2.6		5.9	2.9		--		--		-1.81	.10
M31(Ca-Conf)	6.1	2.3	Career	6.9	2.2		7.0	1.9		-2.37	.05	-2.15	.05	--	--
M39(As-Conf)	6.1	2.2	Assertiveness	5.6	2.4		6.6	1.7		--		--		-2.63	.01
M49	5.8	1.4	Work Ethic	5.3	1.3		5.6	1.3		2.24	.05	--	--	--	--
M51	5.3	1.3	Growth Tension	4.9	1.2		5.0	1.3		2.11	.05	--	--	--	--
V13	41.7	10.6	HDL	44.6	15.6		39.4	11.1		--		--		1.98	.05
V14	5.3	1.9	CHOL/HDL Ratio	4.9	1.5		6.1	3.2		--		--		-2.13	.05
V21	147.3	40.1	LDL	139.9	39.2		155.2	49.3		--		--		-1.72	.10
V22	3.8	1.6	LDL/HDL Ratio	3.4	1.3		4.4	2.4		1.89	.10	--	--	-2.31	.05

2 - tail test

All other differences were not significant at the  $p \leq .10$  level.

TABLE 4

Significant Differences Between Officers with HDL Ratios (CHOL/HDL) at least one-half standard deviation above the mean and those with HDL Ratios at least one-half standard deviations below the mean on predictor and criterion variables.

	Group A HDL Ratio ≥ 6.3		Group B HDL Ratio ≤ 4.3		t	p ≤
	X	SD	X	SD		
	N = 50		N = 78			
P4(E) Dominance	6.6	1.9	7.3	1.9	-1.77	.10
P5(F) Impulsiveness	5.0	2.0	5.7	2.0	-2.11	.05
P7(H) Boldness	5.3	2.2	6.1	2.3	-1.84	.10
P10(M) Imagination	5.7	1.8	6.3	1.8	-1.90	.10
P16(O4) Tension	7.0	1.7	6.4	1.9	1.80	.10
P19(Q1) Extraversion	5.5	2.0	6.1	1.9	-1.81	.10
P20(Q1) Anxiety	6.2	1.6	5.7	1.5	1.80	.10
P22(Q1V) Independence	6.2	1.7	6.8	1.5	-2.21	.05
P24(QV1) Subjectivity	5.3	1.8	5.9	1.8	-1.62	.10
P26 Neuroticism	5.6	1.8	5.0	1.6	2.01	.05
P43 Attn to detail	6.1	2.5	5.1	2.4	2.33	.05
M4(Na-U) Narcism	6.2	1.6	6.7	1.6	-1.71	.10
M7(Ma-U) Mating	4.9	1.6	5.9	1.6	-3.39	.01
M13(Fr-I) Fear	7.0	1.6	6.4	2.1	1.83	.10
M23(Fr-IOT) Fear	6.2	2.8	5.1	2.6	2.30	.05
M24(Na-IOT) Narcism	8.0	2.1	8.7	1.5	-1.93	.05
M25(SE-IOT) Superego	3.7	2.4	3.0	2.0	1.79	.10
M29(AS-IOT) Assertiveness	4.2	2.5	4.9	2.3	-1.65	.10
M37(Ma-CONF) Mating	3.9	2.3	5.1	2.3	-2.94	.01
M49 Work Ethic	5.7	1.2	5.4	1.4	1.66	.10
M50 Sex Confirmation	5.0	1.3	5.7	1.2	-3.44	.01
V11 CHOL	248.3	44.5	183.0	34.9	8.69	.01
V12 TRIGL	139.6	70.4	82.6	32.6	5.34	.01
V13 HDL	32.5	7.5	51.8	14.9	-9.41	.01
V14 CHOL/HDL Ratio	7.9	2.1	3.5	.7	14.21	.01
V16 Uric Acid (UA)	5.6	1.3	5.3	.9	1.72	.10
V17 Sodium(Na)	143.7	3.6	142.6	3.5	1.76	.10
V18 Potassium(K)	4.6	.5	4.4	.4	1.99	.05
V20 ALT	38.7	18.9	29.7	10.2	3.06	.01
V21 LDL	187.9	38.1	114.7	28.9	11.48	.01
V22 LDL/HDL Ratio	6.0	1.5	2.3	.9	15.61	.01

2 - All others were not significant at the  $p \leq .10$  level

TABLE 5

2-tail test  
All other differences were not significant at the  $p \leq .10$  level

TABLE 6

Regression Analysis: Male HDL Ratio Regressed  
on Significant personality, motivational, and  
Biographical variables.

	HDL Ratio		Simple		Multiple		Coef		SIG
	Coefficients		R		R		F Value	DF = 1,202	
Extraversion	-.16		-.21		.21		4.41	<.05	
> one SZ Promotion	1.04		.18		.27		7.69	<.01	
Superego	.11		.18		.31		2.88	<.10	
M25(SE-TOT)	.16		.16		.34		4.22	<.05	
M6(SS-U)	-.83		-.15		.37		5.25	<.05	
Northern Europe	-.12		-.13		.38		2.43	<.15	
Imagination	-1.42		-.05		.40		2.75	<.10	
Slavic	-.11		-.16		.41		3.68	<.10	
Assertiveness	-.10		-.10		.43		2.89	<.10	
Affection									
Constant	7.52								
Multiple R	.43								
R Square	.19								
R Square Shrunk	.15								
Total F Value	4.78								
Significance(DF=9,194)	<.000								
SUMMARIZATION									
	MULTIPLE R		R SQUARE		R SQUARE SHRUNK				
16PF - Personality	.24		.06				.05		
MAT - Motivation	.34		.12				.09		
16PF + MAT	.36		.13				.11		
16PF + DATA + MAT + BIO	.43		.19				.15		

TABLE 7

Regression Analysis: Spouse (female) HDL  
Ratio Regressed on Significant personality,  
motivational, and Biographical variables.

	HDL Ratio		Simple		Multiple		Coef	F Value	DF = 1,107	Sig
	Coefficients	R	R	R	R	R				
P4(E)	.23	.17	.17	.17	.17	.17	4.07			<.01
M3(Fr-U)	-.21	-.13	-.13	.24	.24	.24	2.90			<.05
M19(AS-I)	.19	.16	.16	.27	.27	.27	2.04			<.10
Constant										
Multiple R										
R Square										
R Square Shrunk										
Total F Value										
Significance(DF=3,105)										

## SUMMARIZATION

	MULTIPLE		R SQUARE		R SQUARE	
	R	R	R	R	SHRUNK	SHRUNK
16PE - Personality	.17	.03	.03	.02	.02	.02
MAT - Motivation	.16	.02	.02	.02	.02	.02
16PF + MAT	.27	.07	.07	.05	.05	.05
BIO DATA	NONE	--	--	--	--	--
16PF + MAT + BIO	.27	.07	.07	.05	.05	.05

TABLE 8  
Analysis of variance, for males with CHOL/HDL  
ratio as criterion and cognitive behavior  
patterns as covariates.

	TABP		TBBP	
	A INDEP N=57	B DEP N=46	C INDEP N=44	D DEP N=57
V14 CHOL/HDL Ratio $\bar{X}$ =	4.7	5.4	5.6	5.5
SD =	1.3	2.0	1.7	2.7
Total F Value	2.28			
Significance(DF=3,200)	.08			
Contrasts:				
Group A/B, t Value =	-2.80			
$\underline{p} <$	.01			
Groups A,B/C,D, t value =	-1.77			
$\underline{p} <$	.08			



**1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM**

**Sponsored by the**

**AIR FORCE OFFICE OF SCIENTIFIC RESEARCH**

**Conducted by the**

**SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION**

**FINAL REPORT**

**OPTICAL BISTABILITY WITH LIQUID MEDIA:**

**EXPERIMENTAL STUDIES AND THEORETICAL PREDICTIONS:**

<b>Prepared by:</b>	<b>Dr. Nabil M. Lawandy</b>
<b>Academic Rank:</b>	<b>Assistant Professor</b>
<b>Department and University:</b>	<b>Division of Engineering Brown University</b>
<b>Research Location:</b>	<b>Air Force Geophysics Laboratory Hanscom Airfield</b>
<b>USAF Research:</b>	<b>Dr. D. Katayama</b>
<b>Date:</b>	<b>September 21, 1984</b>
<b>Contract No:</b>	<b>F49620-82-C-0035</b>

OPTICAL BISTABILITY WITH LIQUID MEDIA:  
EXPERIMENTAL STUDIES AND THEORETICAL PREDICTIONS:

by

N. M. Lawandy

and

D. L. MacFarlane and W. S. Rabinovich \*

ABSTRACT

We have experimentally studied optical bistability (OB) in a Fabry-Perot etalon containing a variety of nonlinear liquid media in the quasi steady state limit. We have studied Kerr effect OB using  $CS_2$ , 1,2-dichloroethane, and the effects of the single photon absorber Kodak 9880 dye on this process using Q-switched laser pulses at  $1.05\mu m$ . This work indicates that previous studies using BDN dye dissolved in 1,2-dichloroethane were misinterpreted as the solvent alone can contribute to hysteresis via Kerr effects. In addition we have observed for the first time OB in a liquid using two-photon excitation. These experiments were performed using Rhodamine 6G in methanol solvents. Finally, we have undertaken a theoretical study of the transient response of an adiabatically following, non-dispersive, absorptive medium in a cavity. The results predict an asymptotically stable transient analogous to relaxation oscillations in laser systems. Preliminary experimental results indicate that we may have observed this effect.

---

\* Mr. D. L. MacFarlane and W.S. Rabinovich were graduate students actively participating in this work.

## ACKNOWLEDGEMENT

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for making this research opportunity available to him and his two graduate students. He would also like to express his gratitude to Dr. D. Katayama of the Air Force Geophysics Laboratory for providing us with his facilities, his patience, and the scientific flexibility to undertake new research at AFGL.

In addition, he would like to thank his two graduate students, Mr. D. L. MacFarlane and Mr. W. S. Rabinovich, for their hard work and perseverance. Finally he would like to thank Professor G. S. Heller of Brown University for his generosity and help in using his computing facilities to reduce the stacks of data compiled during our research period.

## I. OBJECTIVES/INTRODUCTION

In recent years the field of optical bistability (OB) has developed dramatically. This explosion in the volume of OB research has resulted from a combination of factors. One of these factors is the underlying goal of achieving an all optical computer. Associated with this theme are questions concerning switching times [1-4], instabilities and temporal behavior [5-8], and tunable hysteresis curves [9]. In addition a great deal of effort has focused on materials studies in the search for low power and fast switching nonlinearities [10-12].

The work we will summarize in this report has dealt with the experimental and theoretical understanding of non-linear Fabry-Perot etalons containing liquid media. Our effort has focused on dispersive switching due to absorbing and non-absorbing media. The absorbing medium studies utilized absorbing dyes in the  $1.06 \mu\text{m}$  region in concentrations of the  $10^{-3}$  to  $10^{-5}$  range. These efforts were aimed at observing single-photon absorptive nonlinearities and two-photon absorptive nonlinearities. The single-photon studies utilized the modelocking dye Kodak 9860 while the two-photon studies utilized rhodamine 6G.

The non-absorbing materials utilized non-linear materials such as  $\text{CS}_2$  and 1,2-dichloroethane. The nonlinearity in  $\text{CS}_2$  is primarily due to the orientational Kerr effect, while electronic nonlinearities are responsible for the effects in 1,2-dichloroethane.

## II. EXPERIMENTAL SECTION

We examined bistability in a 6 mm flat mirror Fabry-Perot etalon containing a 2 mm sample cell. Two different input coupling mirrors were used, one with a transmission of 0.4%, and the other with a 50% transmission. The mirrors were mounted on Burleigh precision laser mounts which allowed for very fine adjustment. The infinite radius of curvature was chosen in order to minimize transverse self-focusing effects in the nonlinear liquids.

The samples were irradiated by  $1.06 \mu\text{m}$  emission from a Molelectron MY-35 Nd:YAG laser at intensities ranging from 20 - 40  $\text{MW}/\text{cm}^2$ . This laser offered consistent 30 nsec. Q-switched pulses.

In all cases the temporal features of the laser pulses were measured using a fast photodiode circuit (Motorola MFO-D1100) constructed in our laboratory. The detector resolved features less than 1 nsec long. The pulses were recorded on a Tektronix programmable digitizer, then drawn by a Hewlett-Packard X-Y plotter controlled by an HP desktop computer. This set up allowed us to average 100 pulses, thereby increasing our signal to noise ratio. The sync out signal of the laser provided a stable external trigger for the digitizer. The experimental arrangement is shown in Figure 1.

In studying the effects of the empty etalon upon the laser pulse we were limited by the fact that the digitizer had only one input channel. This precluded simultaneous measurement of the input and output signals. The consistency of the averaged digitized pulses, however, was high enough to allow for sequential measurements of the input and output pulses. Furthermore this sequential procedure enabled us to use one detector for both measurements; eliminating any error due to the varying electrical characteristics of different detector circuits.

Our early studies of the empty etalon showed instabilities caused by the coupling of the output of the laser back into its gain medium. This problem was eliminated by slightly detuning the front mirror of the etalon. Because of the short length (high Fresnel number) of the etalon and the long path length between the laser and the etalon, the etalon remained in alignment within the diffraction limit.

We first tested for bistability in a number of liquids. Since 1,2-dichloroethane and methanol were to be used as solvents for organic dyes used in the latter part of our studies, it was necessary to determine whether or not they would exhibit bistability. Ethanol and water were also tested.

All the data was analyzed with reference to the incident laser pulse traversing an empty etalon. Calibration checks of the empty etalon exhibited no hysteresis as expected.

In addition we studied  $CS_2$ . This liquid develops a large nonlinear component in its refractive index due to molecular Kerr alignment. This makes it a good candidate for future two color OB experiments.

Finally, we tested two organic dyes. Kodak 9860 in 1,2-dichloroethane was used as a one photon saturable absorber. Rhodamine 6G in methanol was used as a two-photon absorber. These dyes were studied at a variety of concentrations. In all cases, several laser intensities were used.

This data was analyzed at Brown on an Apple II microcomputer using a VeraWriter graphics tablet and customized software developed in our group. Redigitizing with the graphics tablet was necessary to ensure the correct correspondence between the input and output pulses. This correspondence was achieved by starting the redigitizing at identical features on the data curves. Additionally, this put the data in a form amenable to flexible computer analysis. For each experimental case the input and output data sets were normalized and plotted against each other. This allowed us to determine whether hysteresis was present and to examine the time evolution of the hysteresis loop. Figure 2 shows a sample of the unprocessed, digitizer data. The upper curve is the averaged incident input pulse, while the lower curve is the averaged output pulse.

The experimental and theoretical results in the remainder of this report are arranged separately for each experimental effort. The data is presented in raw form as points and a smooth solid curve that approximates a best fit to the points is shown. The results are all presented in terms of the normalized input field intensity  $Y$  and the normalized output field intensity  $X$ .

### III. DISPERSIVE HYSTERESIS BY NON-ABSORBING KERR MEDIA

The studies of our nonlinear etalon using non-absorbing liquids focused on the solvents  $CS_2$ , 1,2-dichloroethane, and methanol. These media may be characterized by an index of refraction  $n=n_0+n_2I$  where  $I$  is the optical intensity in the medium. In addition temporal considerations enter into the problem of driving a nonlinear etalon by an optical pulse.

There are three time constants relevant to the problem. These are: (1) the shortest characteristic time over which the carrier wave must reproduce itself, (2) the input pulse width and (3) the Debye relaxation time of the medium. The Debye relaxation time,  $T_D$ , is the time it takes for an intensity induced index change to decay to its  $e^{-1}$  value once the optical field is removed. The pulse width of the input signal is  $T_P$  and is approximately 30 nanoseconds in our experiments. The time over which the carrier phase must reproduce itself is given by  $T_C=2L/c$  where  $L$  is the etalon mirror separation and  $c$  is the speed of light in the medium. In our experiment,  $T_C$  is on the order of 30 psec. This however is only a first order estimate for Fabry-Perot structures filled with nonlinear media.

Due to the standing wave nature of the field inside a Fabry-Perot etalon, the intensity is a spatially periodic function with period  $\lambda/2$  where  $\lambda$  is the wavelength in the liquid medium.

This results in an index of refraction grating and distributed feedback effects (DFB). Such structures have the effect of altering the optical field time constants and the nonlinear steady state transmission properties of nonlinear etalons. A crude estimate of the time constant for the effects is  $t_c = l/c$  where  $l$  is the path length through the Kerr medium. In our experiments  $l$  is 2mm and  $t_c$  is on the order of 6 psec.

In order to understand the input-output relationship for such media, one must determine whether or not steady state analysis is applicable. The conditions for adiabatic following are given by:

$$T_P \gg T_C \gg T_D \quad (1)$$

This hierarchy assures that the medium does not lag behind in its response to and back onto the field. Previous work by Bischofberger and Shen has shown that even when  $T_C \sim 50T_D$ , hysteresis effects still occur in the input-output curve [13,14].

The limit of a steady adiabatic following may be used to determine the critical field intensities required to observe dispersive switching phenomena and the dependences on losses and mirror reflectivities. Several heuristic partial wave sum methods have been used to analyze OB devices [15,16,17]. These approaches naturally ignore the additional nonlinear reflectivity dependences due to the DFB index grating effects. This is a less serious problem in purely absorptive OB as the absorption grating rapidly approaches zero when the medium's saturation intensity is approached and switching occurs.

More recent work using numerical algorithms to solve the nonlinear wave equation has shown that this approach predicts several branches for the input-output relation which do not actually exist or are unstable [18]. One important point concerning the adiabatic following or steady state limit of the nonlinear Fabry-Perot which has been shown by all methods is that only switch up behavior can occur in input-output curves showing hysteresis. This is the case of an initially non-resonant etalon which, due to the intensity dependent index, becomes more resonant as the input intensity is increased. Since this mechanism has a positive feedback, the system is unstable beyond a certain value of the input intensity and the device exhibits a discontinuous output and jumps to a high value. This is reminiscent of first order phase transitions and can be viewed as such.

The calculation for a dispersive etalon with a mirror reflectivity of 50% is shown in Figure 3. The curve corresponds to a 6mm long etalon filled with  $CS_2$  which is initially in the non-resonant position. This steady state analysis indicates that a value of the order of 146 MW/cm<sup>2</sup> is required to cause switching in such a device. It is important to note that this results in an up switch and counter-clockwise hysteresis.

When the nonlinear etalon is initially resonant, the increasing input field causes the etalon to detune and the internal field to drop. In this arrangement, the feedback is negative and the device does not catastrophically switch. This behavior has been called the power limiter mode in analogy with solid state devices. Therefore in a steady state limit, no clockwise hysteresis can occur.

When the adiabatic regime is no longer accessible, clockwise hysteresis may be observed. This is due to the lag in the medium's response when the field of the input pulse is on its falling transient. Thus the rising part of the pulse produces an index change which does not relax in step with the decaying part of the input pulse and a hysteresis appears in the input-output curves.

### Experimental Results:

The data in this section illustrates the hysteresis effects in Kerr media. The data presented is for  $CS_2$  in an etalon with two 0.4% transmitting mirrors and 1,2,dichloroethane in an etalon with one 0.4% mirror and a 50% transmitting mirror. Figure 4 shows data taken for an empty etalon in order to check the experimental setup. The results clearly show no hysteresis and allow us to determine the degree of scatter in the completely digitized data.

Figures 5a-5c show the input-output results for  $CS_2$  as a function of laser input pulse intensity. Figure 5a shows the effects of finite medium relaxation time and exhibits clockwise hysteresis not observable in a steady state experiment. The laser intensity is  $20MW/cm^2$  and the etalon was initially near resonance. The latter fact can be deduced from the near unity initial slope which occurs in the resonant position. Figures 5b and 5c show multistep (more than one switching event) up switching behavior predicted for an initially non-resonant etalon at intensities of  $30MW/cm^2$  and  $40MW/cm^2$  respectively. The data of Figure 5b shows an up switch, followed by power limiter action, followed by a second up switch. The return path shows the down switch corresponding to the second up switch. Figure 5c shows similar effects as 5b but also indicates that near the peak of the input pulse several hysteresis scenarios occurred. This is due to the presence of multiple peaks in the input pulse (see Figure 2) at high intensities. The device can exhibit clockwise hysteresis between the upper high intensity peaks giving rise to the data points at the top of Figure 5c.

Studies of an etalon filled with solvent 1,2,dichloroethane were also undertaken. This work was done with the expectation that no hysteresis or switching would occur since previous work on BDN dye dissolved in 1,2,dichloroethane attributed hysteresis solely to the BDN molecule [19]. The initial efforts were therefore performed as a check for our subsequent single photon studies of the dye Kodak 9860 dissolved in 1,2,dichloroethane.

The results however, were surprising in that definite switching and hysteresis effects were observed in the pure solvent. Figure 6 shows multiple switching from an initially non-resonant etalon filled with pure 1,2,dichloroethane. The data clearly indicates the expected sharp non-resonant to resonant switching with power limiter action in between. This data has cast serious doubts on the interpretation of the experiments performed on BDN by Zhu and Garmire in reference [19]. It is clear without a doubt that the hysteresis they observed was at the very best due to both BDN and the solvent 1,2,dichloroethane.

### IV. HYSTERESIS WITH KODAK 9860 DYE IN 1,2.DICHLOROETHANE

In this study we were interested in understanding the nonlinear behavior of an etalon containing an absorbing dye. Previous work at  $1.06 \mu m$  using BDN dye in 1,2,dichloroethane observed no hysteresis for the pure solvent and strong hysteresis with BDN solute [19]. This work claims that the counterclockwise hysteresis curves were due to a triplet-triplet transition dispersion effect in the BDN molecule. The range of absorption length products for which they were able to measure hysteresis was  $al \geq 5$ .

Our studies on 1,2,dichloroethane discussed in the preceding section clearly reveal that this solvent can act as a Kerr medium and generate up switching hysteresis alone. The addition of Kodak 9860 to 1,2,dichloroethane at a variety of concentrations in the etalon with one 50% transmitting mirror and one 0.4% transmitting mirror yielded more pronounced switching behavior. Figures 7a and 7b show the hysteresis observed for our etalon with a  $1.06 \mu m$  absorption length product of  $al=1.3$ . Figure 7a shows the response of an initially non-resonant etalon which switches up when subjected to  $20MW/cm^2$  input pulses. Figure 7b shows the response of

our etalon with the same absorption length product to  $30\text{ MW/cm}^2$  input pulses. This case corresponds to an initially resonant etalon and shows the power limiter action after the first knee.

Figure 7c shows the response of the etalon with an  $\alpha l = 0.33$  and  $20\text{ MW/cm}^2$  input pulses. The results show examples of sharp upper branch down switch. Figure 7d is the same concentration as 7c but with  $30\text{ MW/cm}^2$  input pulses.

Our studies show that the addition of a single-photon absorber to the already nonlinear solvent 1,2-dichloroethane widens and alters the solvent hysteresis behavior. The effect of the absorber could be to contribute additional nonlinear dispersion. This, however, is not as likely since this effect, due to the resonant nature of the transition, could add to or subtract from the Kerr nonlinearity of the 1,2-dichloroethane depending on whether the laser frequency was higher or lower than the dye absorption band. Since the hysteresis occurs for both Kodak 9860 in our work and BDN in reference [19], we suspect that this is not the effect at work.

Theoretical studies of our etalon with a Kerr nonlinearity and a saturable absorber have been undertaken in our group. A simple model yields the following input-output relation in steady state:

$$(1-R)^2 Y = X \left[ e^{+2p} + R^2 e^{-2p} + 2R \cos(2k_0 n_2 l_p X) \right] \quad (2)$$

where  $p = \frac{\alpha_p l}{1+X}$  and X and Y are now in units of  $I_s$ , the saturation intensity of the absorber;  $n_2$  is the nonlinear Kerr coefficient of the solvent and  $k_0$  is the wavevector at low intensity. Analysis of Equation 2 has shown that  $\alpha_p l$  values on the order of 1 - 2.5 cause the hysteresis curves of the solvent alone ( $\alpha_p l = 0$ ) to widen and develop more defined features. This is due to the effect of the saturable absorber. It should be noted that the absorber alone ( $n_2 = 0$ ) does not produce hysteresis. A comparison of  $\alpha_p l = 0$  and  $\alpha_p l = 2.4$  results in a five-fold increase in the width of the pure Kerr effect hysteresis.

We therefore conclude that the most likely explanation of the results in reference [19] on BDN and our work with Kodak 9860 in 1,2-dichloroethane is that the solvent is a weak Kerr medium and the dye acts as a partially saturable absorber. The dye dispersion as we have shown is not necessary to produce wider hysteresis behavior if the solvent is already a Kerr medium. Finally, we must conclude that no single-photon absorber studies to date have shown optical switching and/or hysteresis solely due to the added dye molecule.

#### V. HYSTERESIS WITH TWO-PHOTON NONLINEAR DISPERSION IN RHODAMINE-METHANOL SOLUTIONS

We have performed an experimental study of the nonlinear behavior of an etalon containing rhodamine 6G in methanol with input pulses at  $1.06\mu\text{m}$ . Since rhodamine absorbs in the region centered around  $5300\text{ \AA}$ , any response due to the dye must occur via a two-photon mechanism.

In order to remove the uncertainty associated with combined dye solvent effects, we utilized methanol and tested the response of the 50% transmitting etalon filled only with solvent. The results, shown in Figure 8, indicated no Kerr effects in methanol. This experiment allows us to attribute all nonlinear behavior with rhodamine 6G to the two-photon process.

Experiments were performed in the 50% transmitting etalon with a variety of concentrations. Since the two-photon absorption mechanism does not follow Beer's law, we can only specify the molar concentrations of each experiment. Figure 9a shows the response of a  $2.0 \times 10^{-4}\text{ M}$  solution to  $20\text{ MW/cm}^2$  input pulses. The data shows a narrow hysteresis curve. When



the same input intensity is utilized but the concentration is increased to  $2.4 \times 10^{-4}$  M, sharp well-defined multiple up switching is observed. This is shown in Figure 9b. Figure 9c shows the hysteresis which occurs when the higher concentration ( $2.4 \times 10^{-4}$  M) is subjected to 40 MW/cm<sup>2</sup> input pulses. This etalon shows a power limiter type behavior and an extremely wide hysteresis.

The work with rhodamine 6G is novel in two ways. It is the first nonlinear etalon in which the hysteresis and switching can be clearly attributed to the dye molecules, and the first nonlinear switching due to a two-photon mechanism.

We are at present unsure as to whether the effects observed are primarily due to the dispersive or absorptive component of the two-photon process. Our inclination at this point is to attribute the effects to dispersion as the onset for the two-photon saturated absorption process in rhodamine is in the 0.7 - 1.0 GW/cm<sup>2</sup> range [20]. Further theoretical and experimental studies of this new effect are underway with Dr. Kanyama in his laboratory at AFGL.

## VI. OPTICAL TRANSIENTS IN THE NONLINEAR ETALON

The dynamics of optically bistable systems has been a subject of much interest in recent years. This interest extends from the practical questions concerning switching times for all-optical computations to the esoteric questions of stability, the transition to chaos, and optical turbulence. The questions of switching times in the bad cavity limit and good cavity limit have been investigated using mean-field models [21] and the saturable interferometer approach [17] for absorbing two level systems. In addition, Bonifacio and Lugiato have predicted a purely absorptive instability at high fields [6].

The question of pulsations and oscillation in dispersive OB have been studied by Ikeda [22] and others. Ikeda's work has predicted the Ikeda instability and period doubling bifurcations to optical chaos in a ring geometry. Silberberg and Bar-Joseph have shown that a four-wave mixing instability leading to chaos may occur in a system where two unrelated optical fields counterpropagate through a Kerr medium [23]. Goldstone and Garmire have predicted a regenerative oscillation regime for Fabry-Perot etalons containing Kerr media [24].

We have experimentally examined the output of our etalon filled with 1,2-dichloroethane solution containing the absorbing dye Kodak 9860 for transients. In a narrow concentration range we have observed oscillatory behavior on the down slope of the input pulse. Figures 10a and 10b show the average of 100 input and output pulses respectively. The etalon used had two 99.6% reflecting mirrors and the absorption length product for the dye at  $1.06 \mu\text{m}$  was 1.5. Figure 10b clearly exhibits oscillatory structure with a 2-3 nsec period.

The regenerative pulsations predicted by Goldstone and Garmire have oscillation periods of  $5-7 T_R$ . Using our 6mm etalon, this predicts periods in the 200 psec range. This value is off by a factor of 100 and is therefore indicative that the dispersive effects are not responsible.

We propose an explanation for these transients based on the coupling between the population difference in the absorbing dye and the cavity photon field, similar to the relaxation oscillations encountered in inverted laser systems [25].

Since the interaction of the cavity field with the population difference in an absorbing system does not have a restoring force type of coupling (as in an inverted laser transition), this effect would require driving the system by an input of negative slope or one which decreases with time. This is completely consistent with the observation of these transients on the down slope of the input pulse.

The basic equations which describe the system are given below:

$$\dot{\Delta} = -[(N + \Delta)\Gamma + 2BI\Delta] \quad (3a)$$

$$\dot{I} = \frac{I_0(t)}{T} + I(\sigma \cdot \Delta - \frac{1}{t_c}) \quad (3b)$$

$\Delta$  is the difference between the excited singlet state ( $S_1$ ) and ground singlet state ( $S_0$ ) populations,  $N$  is the total population, and  $T$  is the mirror transmission.  $\Gamma$  is the rate of relaxation from  $S_1$  to  $S_0$ ,  $B$  is Einstein coefficient and  $t_c$  is the cavity lifetime (not to be confused with  $T_c$ , the roundtrip time),  $I$  is the internal average intensity and  $I_0(t)$  is the time derivative of the input pulse. We assume in all calculations a triangular pulse with input pulse rise  $I_0(t) = d_1 t$  and decay  $I(t) = I_0(1 - d_2 t)$ .

These equations have been solved for the steady state values and a perturbation stability analysis was performed around these values. The results yield a damped oscillation about steady state indicating that the system is asymptotically stable [25]. The expression for the oscillation frequency,  $\Omega$ , is given by:

$$\Omega^2 = \frac{\Gamma}{t_c} - \Delta_{ss}\Gamma\sigma c + \frac{2I_{ss}B}{t_c}$$

where  $\Delta_{ss} = \frac{-N\Gamma}{\Gamma + 2BI_{ss}}$  and  $I_{ss}$  satisfies:

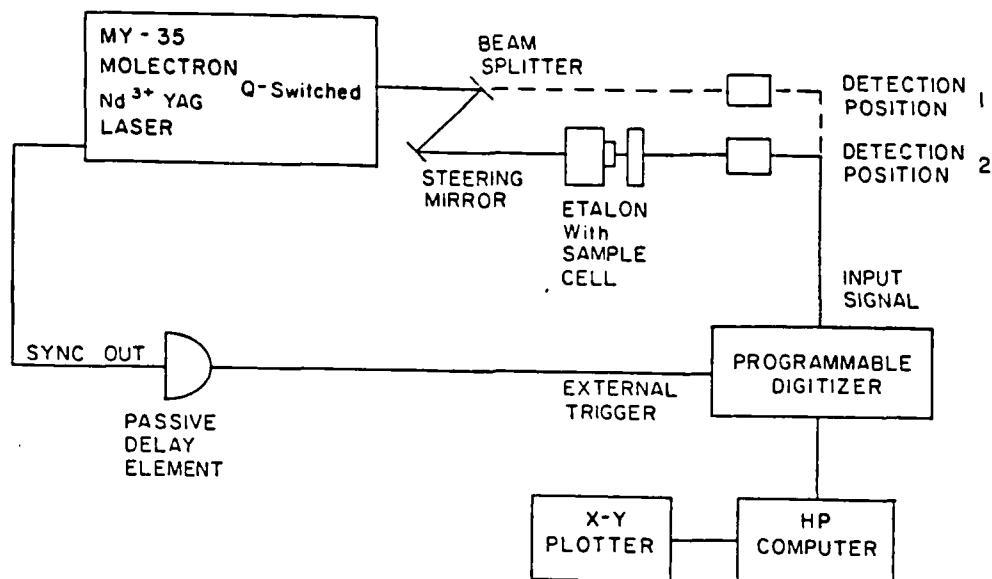
$$I_{ss} + I_{ss} \left[ -\frac{d_2 t_c}{T} + \frac{N\Gamma\sigma c t_c}{2B} + \frac{\Gamma}{2B} \right] - \frac{d_2 \Gamma t_c}{2BT} = 0$$

The values for Kodak 9860 [26] result in our angular frequency of  $3 \times 10^9$  radians/sec. This is equivalent to a 2 nsec period and is in excellent agreement with the experimental observations.

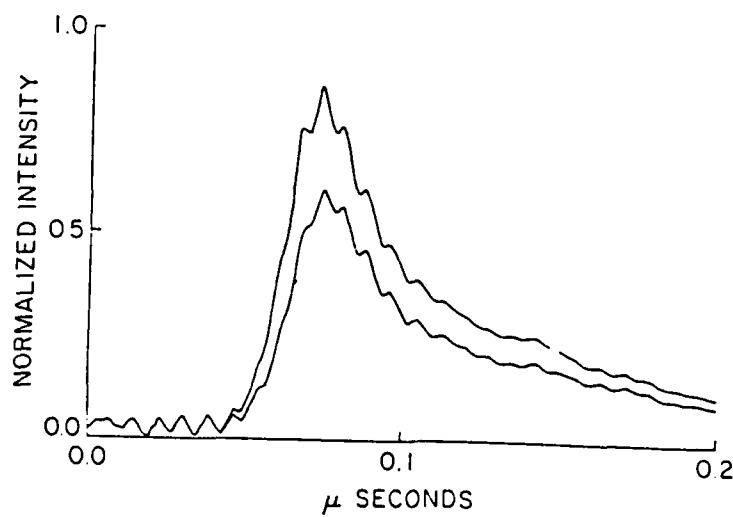
## VII. RECOMMENDATIONS

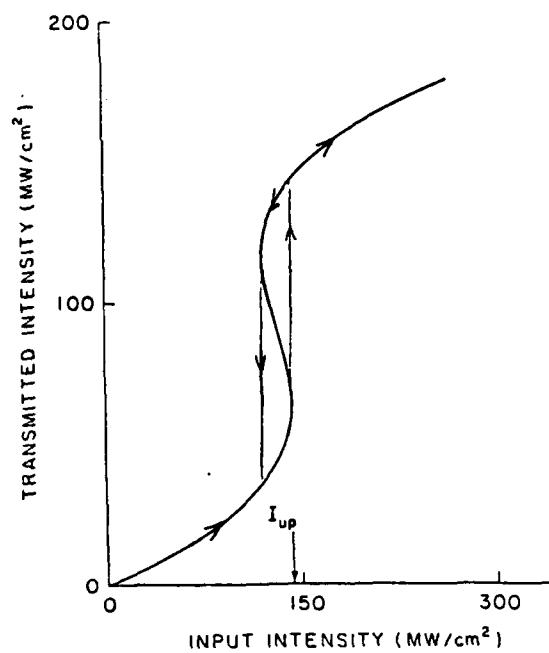
The work we have performed has resulted in increased understanding of nonlinear etalons containing absorbing media. We have shown that previous work using 1,2-dichloroethane was misinterpreted and have proposed a theory for the effects of including saturable single-photon absorption with a Kerr nonlinearity. Our work has also predicted an oscillation transient for such systems. The two-photon studies are the first of their kind and have demonstrated the first optical hysteresis effects due to a solute molecule.

Based on this summer's accomplishments, we recommend that follow on work be continued along three fronts: (1) study the dynamics of two-photon dispersion and switching for possible two-photon laser modelocking applications, (2) continue experimental work on transient oscillations in absorptive and dispersive Fabry-Perot etalons and include coherent effects in the theoretical treatment, and (3) develop a two color optically bistable etalon device using Kerr media and stimulated Raman parametric oscillation. The theory and preliminary experimentation has been prepared this summer for a fruitful effort in all of the above areas.

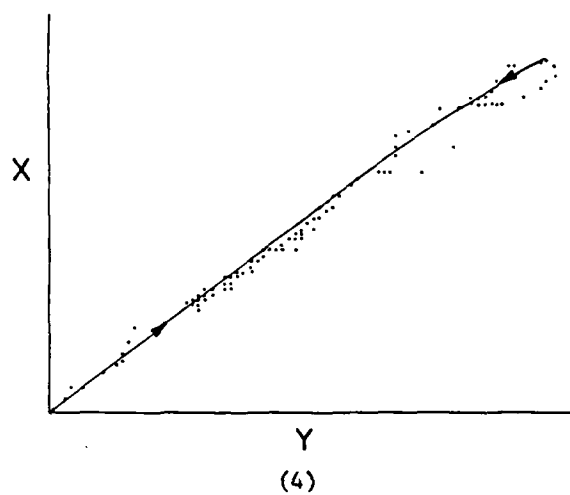


(1)

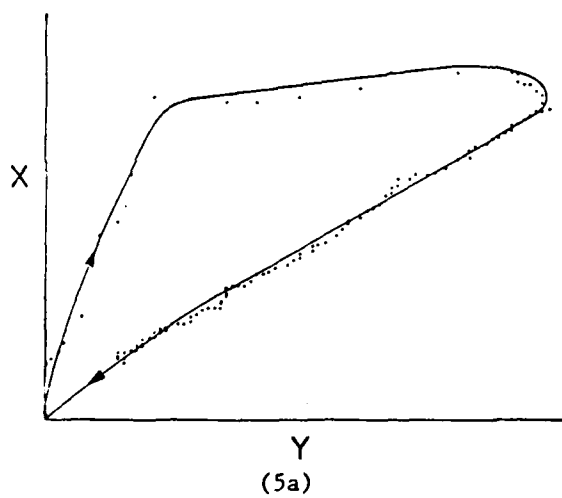




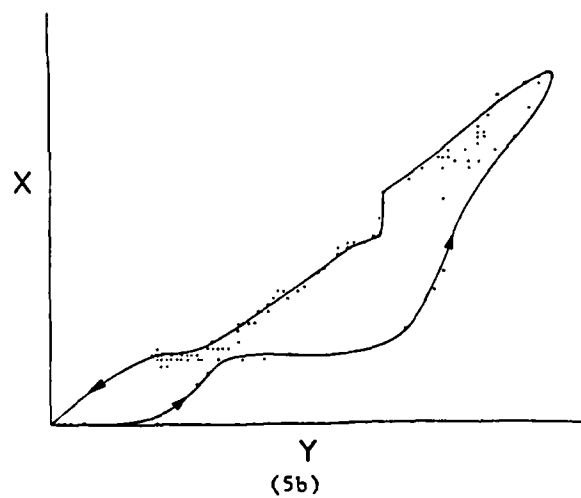
(3)



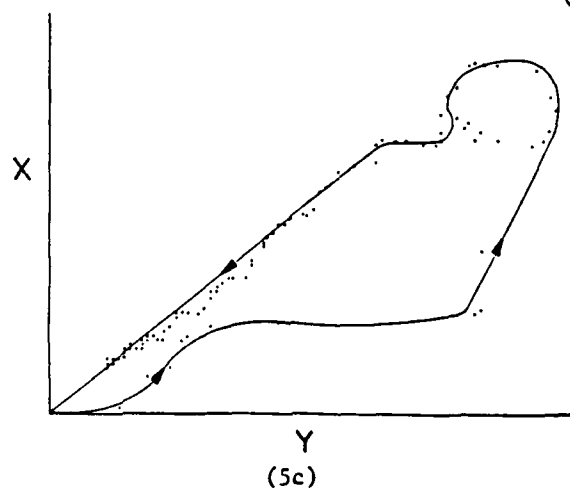
(4)



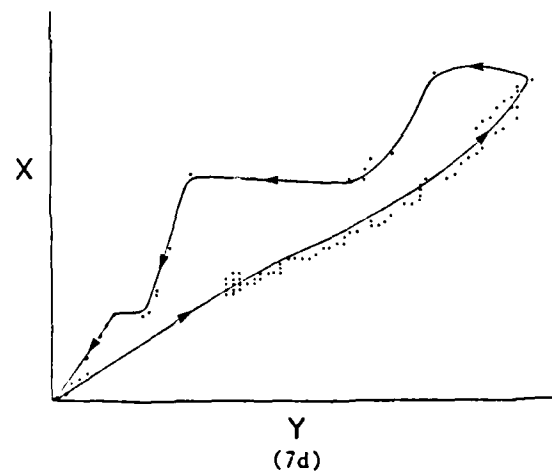
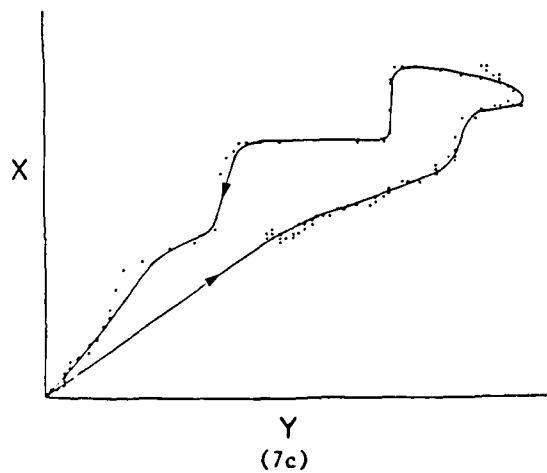
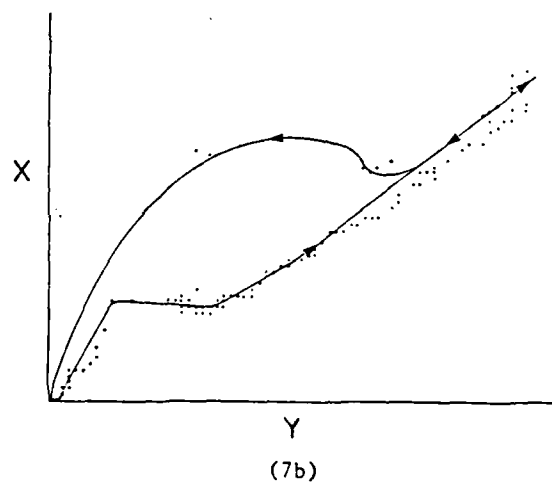
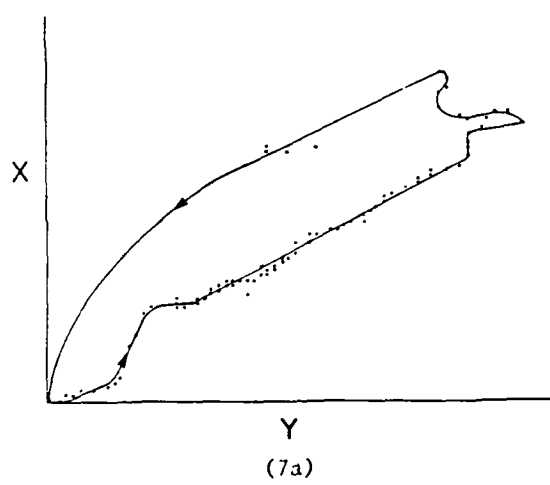
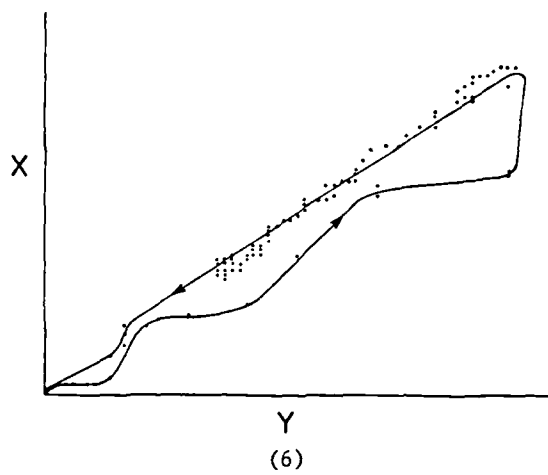
(5a)

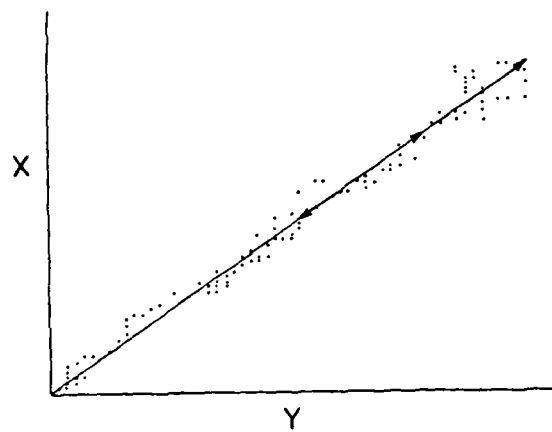


(5b)

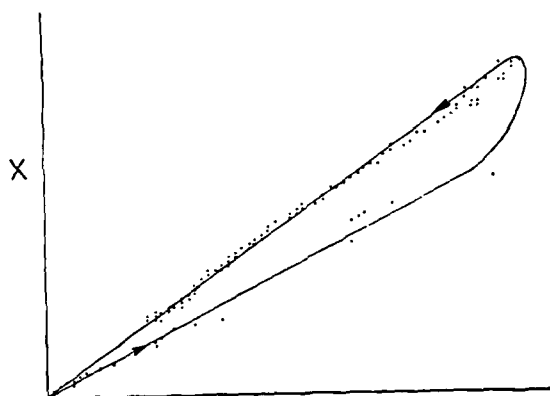


(5c)

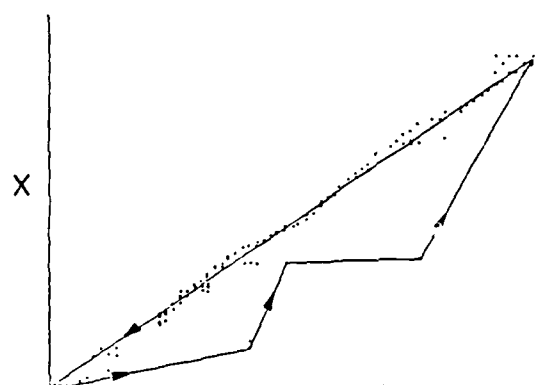




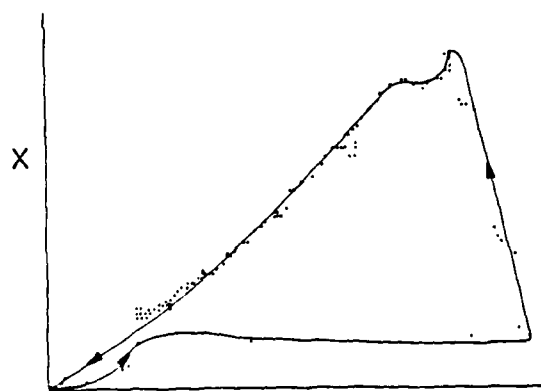
(8)



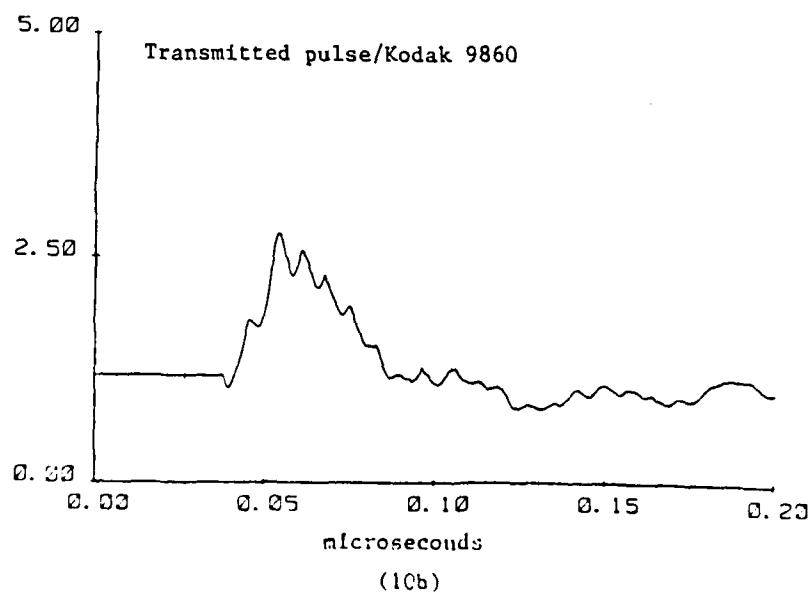
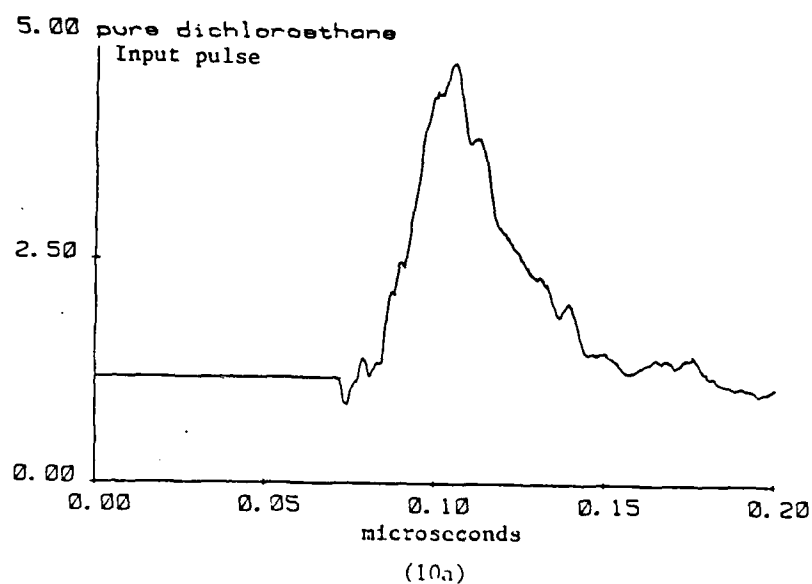
(9a)



(9b)



(9c)



## REFERENCES

1. R. Bonifacio and P. Meystre, "Transient response in optical bistability," Opt. Commun., vol. 27, pp. 147-150, 1978.
2. V. Benza and L. A. Lugiato, "Analytical treatment of the transient in absorptive optical bistability," Lett. Nuovo Cimento, vol. 26, pp. 405-409, 1979.
3. P. Mandel and T. Erneux, "Dynamics of nascent hysteresis in optical bistability", Opt. Commun., vol. 44, pp. 55-58, 1982.
4. J. V. Maloney and H. M. Gibbs, "Role of diffractive coupling and self focusing or defocusing in the dynamic switching of a bistable optical cavity" Phys. Rev. Lett., vol. 48 (23), pp. 1607-1610, (1982).
5. R. Bonifacio, M. Gronchi, and L. A. Lugiato, "Self-pulsing in bistable absorption," Opt. Commun., vol. 30, pp. 129-133, 1979.
6. R. Bonifacio and L. A. Lugiato, "Instabilities for a coherently driven absorber in a ring cavity," Lett. Nuovo Cimento, vol. 21, pp. 510-516, 1978.
7. M. Gronchi, V. Benza, L. A. Lugiato, P. Meystre, and M. Sargent, III, "Analysis of self-pulsing in absorptive optical bistability," Phys. Rev. A, vol. 24, pp. 1419-1435, 1981.
8. M. W. Derstine, H. M. Gibbs, F. A. Hopf, and D. L. Kaplan, "Alternate paths to chaos in optical bistability," Phys. Rev. A, vol. 27, pp. 3200-3208, 1983.
9. N. M. Lawandy, and W. S. Rabinovich, "Absorptive bistability in a three-level system interacting with two fields" IEEE J. Quant. Electron vol. QE-20, No. 5, May 1984.
10. N. M. Lawandy, D. V. Plant and D. L. MacFarlane, "Optical bistability in an expanding thermally dissipative etalon," Submitted to IEEE J. Quant. Electron, 1984.
11. D. A. B. Miller, S. D. Smith and A. Johnston, "Optical bistability and signal amplification in a semi-conductor crystal: Applications of new low-power nonlinear effects in InSb", Appl. Phys. Lett., vol. 35, pp. 658-660, (1979).
12. D. Grischkowsky, "Nonlinear Fabry-Perot interferometer with sub-nanosecond response times," J. Opt. Soc. Amer., vol. 68, p. 641, (1978).
13. T. Bischofberger and Y. R. Shen, "Transient behavior of a nonlinear Fabry-Perot", Appl. Phys. Lett., vol. 32, pp. 156-158, (1978).



14. T. Bischofberger and Y. R. Shen, "Theoretical and experimental study of the dynamic behavior of a nonlinear Fabry-Perot interferometer", Phys. Rev. A, vol. 19, pp 1169-1176, (1979).
15. F. S. Felber and J. H. Marburger, "Theory of non-resonant multistable optical devices," Appl. Phys. Lett., vol. 28, pp. 731-733, 1976.
16. D. A. B. Miller, "Refractive Fabry-Perot bistability with linear absorption: Theory of operation and cavity optimization", IEEE J. Quant. Electron QE-17 (3), p. 306-311, 1981.
17. N. M. Lawandy and R. P. Willner, "Saturable interferometer theory of optical bistability", IEEE J. Quant. Electron, QE-19 (10), p. 1499-1505.
18. Y. B. Band, "Theory of optical bistability in nonlinear media for slowly ranging incident intensity", Optical Bistability 2, Ed., C. M. Bowden, H. M. Gibbs and S. L. McCall, Plenum Press, New York, 1983, p. 463-469.
19. Z. F. Zhu and E. M. Garmire, "Optical bistability in BDN dye", IEEE J. Quant. Electron, QE-19 (9), p. 1495-1498, 1983.
20. J. P. Hermann and J. Ducuing, "Dispersion of the two-photon cross section in rhodamine dyes", Opt. Comm. 6 (2), p. 101-105, 1972.
21. P. Mandel and T. Erneux "Switching times in absorptive optical bistability", Opt. Comm. 42 (5), p. 362-365, 1982.
22. K. Ikeda, "Multiple valued stationary state and its instability of the transmitted light in a ring cavity system", Opt. Comm., vol. 30 (2), pp. 257-261, 1979.
23. I. Bar-Joseph and Y. Silberberg, "The mechanism of instabilities in an optical cavity", Opt. Comm., vol. 48 (1), pp. 53-56, 1983.
24. J. A. Goldstone and E. M. Garmire, "Regenerative oscillation in the nonlinear Fabry-Perot interferometer", IEEE J. Quant. Electron., vol. QE-19 (2), pp. 208-217, 1983.
25. N. M. Lawandy, "Relaxation oscillations and instability", IEEE J. Quant. Electron, vol. QE-18 (12), pp. 1992-1994, 1982.
26. A. Seilmeier, B. Kopainsky, W. Kaiser, "Infrared fluorescence and laser action of fast mode-locking dyes", Appl. Phys. 22, pp. 355-359, 1980.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

SURVEY OF HIGH-ENERGY MOLECULAR SYSTEMS

Prepared by:	Dr. E. Miller Layton, Jr.
Academic Rank:	Associate Professor
Department and School:	Chemistry/Applied Mathematics Sterling College, Sterling, Kansas
Research Location:	Air Force Rocket Propulsion Laboratory, Liquid Rocket Division, Chemistry and Materials Branch, Basic Chemical Research Section.
USAF Research:	Dr. Stephen L. Rodgers
Date:	September 15, 1984
Contract No:	F49620-82-C-0035

# SURVEY OF HIGH-ENERGY MOLECULAR SYSTEMS

by

E. Miller Layton, Jr.

## ABSTRACT

We have undertaken a theoretical survey of unusual molecules with potential for storing large amounts of energy in a conveniently accessible form. We have especially considered molecular systems containing helium because of the large energy separation between ground and excited states in the helium atom and also because of recent work indicating the possibility of an unusual stabilizing mechanism for helium in the high-spin triplet state. The stabilizing mechanism is identified, promising systems are catalogued, theoretical calculations are proposed, and possible synthetic pathways are reported.

#### ACKNOWLEDGMENTS

Preliminary significant discussions of ideas and approaches were held with Professor Donald W. Setser of Kansas State University and Dr. Ronald Herm of the Aerospace Corporation. A continuing profitable dialogue was maintained with Dr. Stephen L. Rodgers of the Air Force Rocket Propulsion Laboratory during the assignment. The author is especially grateful to Mr. John T. Nakamura, LKL Branch Chief at the Rocket Propulsion Laboratory, for the gracious hospitality afforded him during this appointment at the Air Force Rocket Propulsion Laboratory. The author would also like to thank First Lieutenant William Sowell, U.S. Air Force, for several significant contributions to this work and First Lieutenant Michael B. Coolidge for numerous instances of courteous assistance. This work was supported by the Air Force Systems Command, and the U.S. Air Force Office of Scientific Research through their Summer Faculty Research Program.

## I. INTRODUCTION

The interest of the Air Force in stable systems of high energy is obvious and well-known. Applications abound in the areas of conventional and rocket fuels, explosives and laser systems. In recent years a new class of systems has engendered interest because of its novelty, very high energy potential, and critical test relationship to the de Boer quantum theorem of corresponding states (QTCS). This class of systems is that of spin-polarized atoms, meaning atoms with all valence electron spins aligned. Since the prediction of existence of metallic hydrogen (1), much work has been done on spin-polarized atomic hydrogen (H $\uparrow$ ) (see 2 for representative references). A system deriving from spin-polarized helium (He $\uparrow$ ) has been developed over the last decade (2) and recently a report on spin-polarized nitrogen (N $\uparrow$ ) has appeared in the literature (3). The interest in these spin-polarized systems relates to their theoretically expected collapse into stable, high-energy, spin-aligned condensed phases which may then be employed as fuel reservoirs for various applications (2,3). The reference by Ferrante and Stwalley (3) has a particularly lucid discussion of the de Boer QTCS (4).

The promise of these high-energy systems has led us to seek other molecular systems which would have the desirable attributes of high free-energy content and kinetic stability against spontaneous decay. Reference to Figure 1 illustrates that atomic helium is particularly attractive since the metastable triplet level ( $^3S_1$ ) lies 19.8 eV (or 457.0 kcal/mole) above the ground state. If helium could be stored in this metastable state by molecular formation, then the decay lifetime might be increased beyond the  $2\frac{1}{2}$  hours recognized for the isolated, metastable helium atom in state T\* (5).

## II. OBJECTIVES OF THE RESEARCH EFFORT

This project was undertaken to survey other potentially high-energy containing systems, both from a theoretical and an experimental viewpoint. Theoretically, our aim has been to inspect the electronic structures of promising molecular systems as a function of geometrical parameters, using accepted theoretical approaches, and to identify

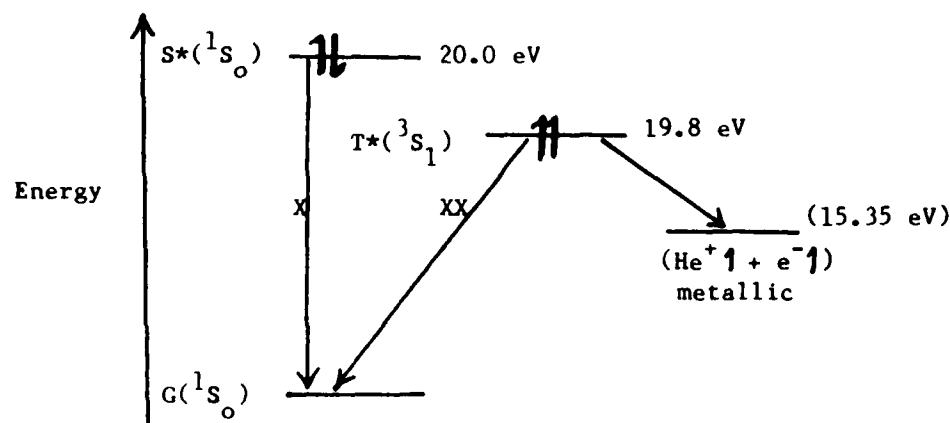


FIGURE 1. SCHEMATIC REPRESENTATION OF HELIUM ENERGETICS. Arrows on the energy levels indicate total spin properties of occupying electrons:  $\uparrow\downarrow$ , paired or singlet;  $\uparrow\uparrow$ , unpaired or triplet. Both excited states ( $S^*$ , singlet, and  $T^*$ , triplet), the lowest in their respective manifolds, are metastable with respect to combination with the ground state. For the excited singlet state, coupling with the ground state is forbidden because of the lack of change in angular momentum during the transition. For the excited triplet state, combination with ground is doubly forbidden by reason of angular momentum conservation and the necessity of a multiplicity (spin) change. X's on transition arrows indicate the extent of the forbidden character. The position of the predicted metallic state of spin-polarized helium is indicated at a lower energy than either atomic state, hence presumably further stabilized with respect to rapid decay to ground.

those systems which have bound electronic states, ground or excited, neutral or ionic. This theoretical identification of bound molecular states should (i) guide the synthetic chemist along paths towards systems which may be isolable and stable and (ii) emphasize systems which may be further considered for condensation by condensed state theoreticians, either along the lines of the de Boer theory or by other mechanistic pathways. The end product would be energetic material in a stable bulk form, which could provide energy according to the demands of an external control system.

### III. PROSPECTIVE MOLECULAR SYSTEMS

We have attempted a systematic consideration of molecules which show promise of stability in high energy configurations. Many of our

molecular systems include helium atoms since a bonding situation with helium should invoke a significant fraction of excited state ( $S^*$  or  $T^*$ ) character. Such a system, even as a ground state molecular system, should contain a large admixture of excited state energy. Electronic properties of molecular systems presented in the following tables may be calculated by standard prescriptions (see 6 or 7)\*. Literature references to theoretical structure calculations (9) and experimental data (10) are summarized for a limited number of diatomic molecules. For helium molecular systems in general, however, polyatomic as well as diatomic, experimental references prove to be surprisingly sparse. Companion volumes on theoretical structural calculations (11) and experimental references (12) are available for polyatomic systems, but no molecular systems larger than diatomic are experimentally catalogued. Since numerous stable spectroscopic states are reasonably well-known, it is quite surprising that more investigations have not been reported. This may reflect a lack of interest deriving from laboratory synthetic difficulties, or a lack of an important goal for stabilizing synthetic products. The forms of combining orbitals are first approximations, normally helium  $1s$  orbitals being employed. In actual fact, bonding may very well require  $2s$  orbitals on helium, with the same symmetry as  $1s$ , or helium  $2p$  orbitals, with an additional systematic complication in the bonding description. These refinements automatically appear in the computer results of variational calculations of electronic structures of the molecules as long as provision is made to allow such higher energy orbitals to be invoked.

Table I presents the important features of primitive helium systems (properties of the atom in the ground and lowest excited states) and Figure 2 establishes our energy and spin conventions. More sophisticated wave functions are available for helium (its total energy has been calculated more accurately than it can be measured (13)), but for comparison we select conceptually simpler representations, which do not

\*For instructional purposes it should be mentioned that the methods of second quantization are carefully and thoroughly developed for the  $\text{HeH}^+$  system as problems at the end of each chapter in a recent reference (8).

SYSTEM	ELECTRONIC STRUCTURE	NON-NORMALIZED WAVE FUNCTION	COMMENTS
He	$1s^2$	$1s(1)1s(2)[\alpha(1)\beta(2) - \beta(1)\alpha(2)]$	Ground state of atomic helium, $G = {}^1S_0$ , precursor or any molecular helium system.
He*	$1s2s$	$[1s(1)2s(2) + 2s(1)1s(2)]$ $\times [\alpha(1)\beta(2) - \beta(1)\alpha(2)]$	Excited singlet atomic helium state, $S^* = {}^1S_0$ . (Valence bond description)
		$[1s(1)2s(2) - 2s(1)1s(2)]$ $\times \begin{cases} \alpha(1)\alpha(2) \\ \alpha(1)\beta(2) + \beta(1)\alpha(2) \\ \beta(1)\beta(2) \end{cases}$	Excited triplet atomic helium state, $T^* = {}^3S_1$ . (Valence bond description)

TABLE I. THEORETICAL FEATURES OF PRIMITIVE HELIUM SYSTEMS. Note that  $1s(1)$  refers to an atomic orbital (AO) on the helium atom, containing electron 1.  $\alpha$  and  $\beta$  refer to electron spins,  $\uparrow$  and  $\downarrow$ . See reference 6, section 4.3 and 4.4, for instance, for a careful discussion of AO's. Introductory atomic and molecular structure theory are clearly presented in the same reference, chapters 9, 10, 11 and 12.

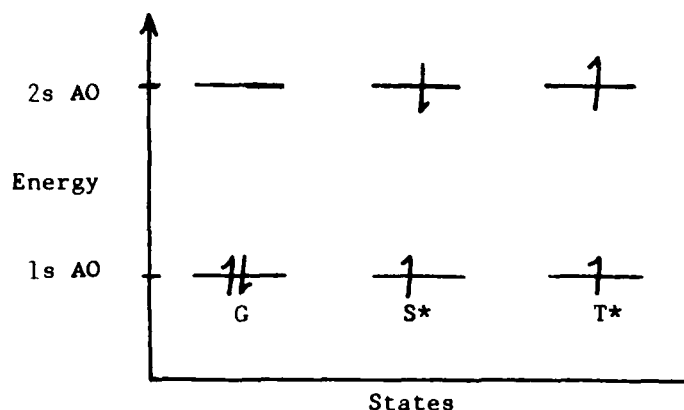


FIGURE 2. ORBITAL OCCUPATION OF LOWEST ENERGY STATES OF HELIUM. Energy increases vertically. Arrows indicate spin of particular electron. Following the Pauli Exclusion Principle, two electrons in the same energy level must have opposed spins.



alter the model for understanding the structure of helium. Tables II - VI display a selection of molecules, with proposed electronic structures, symmetry classifications and conceptual bonding schemes. Note that the geometric form of the orbitals is developed in these tables rather than recourse to analytic expressions. Analytic wave function expressions are available (14), but their interpretation becomes progressively more difficult with the addition of each electron. Atomic orbitals (AO's) are characteristically combined to form molecular orbitals (MO's).

Table II presents a selection of the simplest helium hydride systems which we considered. Each added electron requires an additional factor in the analytical wave function expression. Spin is formally neglected in these wave functions, but incorporated in the electron occupation diagram by opposed arrows.  $\text{HeH}^{++}$  is the simplest heteronuclear diatomic molecular system and plays the same rôle that  $\text{H}_2^+$  plays in homonuclear diatomic molecular theory. The ground state of  $\text{HeH}^{++}$  is unstable and separates into  $\text{He}^+$  and  $\text{H}^+$  ions. The  $\text{HeH}^+$  system is stable and excitation dissociates it into  $\text{He}$  and  $\text{H}^+$ . This system displays a simple MO wave function, characterized by delocalization, or association of each electron with both nuclear attraction centers. The last of these systems,  $\text{HeH}$ , is also not experimentally stable. As isolated systems, these hydrides do not show promise as high energy storage species.

Table III presents several geometries of more complicated hydrides.  $\text{HeH}_2$  affords an initially attractive possibility for a stable system since a central helium atom in a triplet state can be considered as attractive to two hydrogen atoms with spins that can pair with the unpaired helium electrons. Unfortunately, there is no preferred directional character to the resultant bonds and, of rather more importance, the deeper energy of the helium atomic system accepts a hydrogen electron in its 1s shell, then repels both additional electrons to form an excited separated  $\text{H}_2$  molecule (see figure 3). The bent form of  $\text{HeH}_2$  represents an isolated helium atom approaching a hydrogen molecule in an unsymmetric arrangement. Calculations (15) indicate a minimum in the potential energy

SYSTEM	ELECTRONIC STRUCTURE (STATE DESIGNATION)		FORM OF WAVE FUNCTION
HeH <sup>++</sup>	$1\sigma_g$	$(^2\Sigma^+)$	$c_1 1s_{\text{He}}(1) + c_2 1s_{\text{H}}(1)$
HeH <sup>+</sup>	$1\sigma_g^2$	$(^1\Sigma^+)$	$[c_1 1s_{\text{He}}(1) + c_2 1s_{\text{H}}(1)]$ X $[c_1 1s_{\text{He}}(2) + c_2 1s_{\text{H}}(1)]$
HeH	$1\sigma_g^2 1\sigma_u$	$(^2\Sigma^+)$	$[c_1 1s_{\text{He}}(1) + c_2 1s_{\text{H}}(2)]$ X [same function for electron 2] X [same function for electron 3]

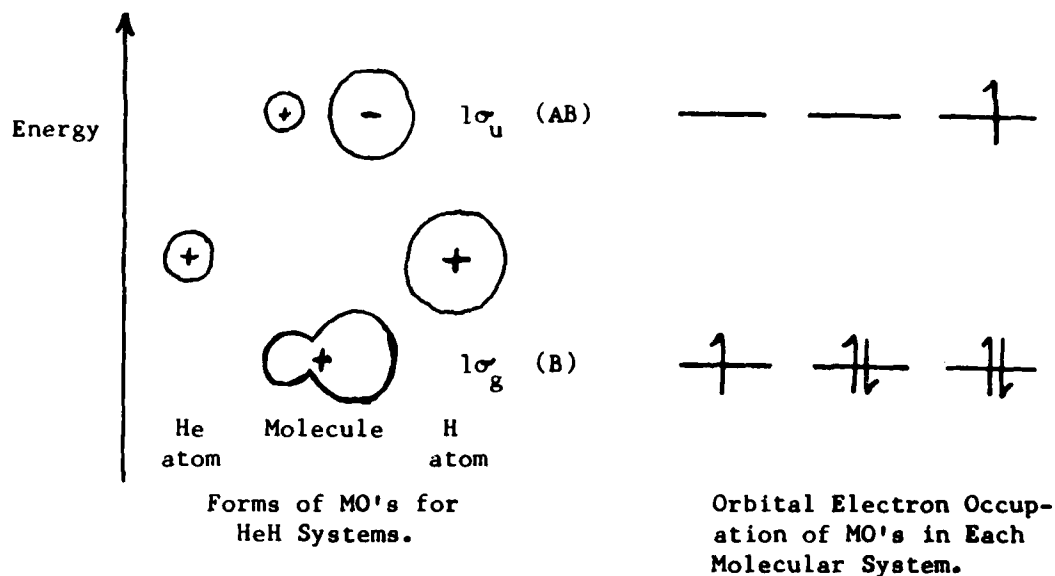


TABLE II. DIATOMIC HELIUM HYDRIDE SYSTEMS. Note that  $1s_{\text{He}}(1)$  specifies that electron 1 is in a  $1s$  AO on the helium atom. The algebraic sign indicates the positive or negative value of the wave function in space, not the sign of the charges in the orbitals. The  $c_i$ 's in the wave function expressions are mixing coefficients which determine the fractional contribution of each AO. The lower part of the table schematically indicates the relative energy level positions, the geometric form of the atomic and molecular orbitals and the electronic populations of the levels in the particular molecular system. Bonding (B) and antibonding (AB) character are identified with particular MO's and energy levels.

SYSTEM	FORM OF WAVE FUNCTION			ORBITAL POPULATIONS	
$\text{HeH}_2$ (linear) ( $D_{\infty h}$ )	Energy ↑		AB		
			NB		
			B		
$\text{HeH}_2$ (bent)		 ( $C_{2v}$ )			
		 ( $C_s$ )			
$\text{He}_2\text{H}$ (linear) ( $D_{\infty h}$ )	Energy ↑		AB		
			NB		
			B		

TABLE III. TRIATOMIC HELIUM HYDRIDES. Diagrams indicate the various  $1s$  AO's which form MO's. The orbital populations then refer to the signified MO's. Bonding character (B), non-bonding character (NB) and antibonding or repulsive character (AB) are identified with particular MO's and energy levels.

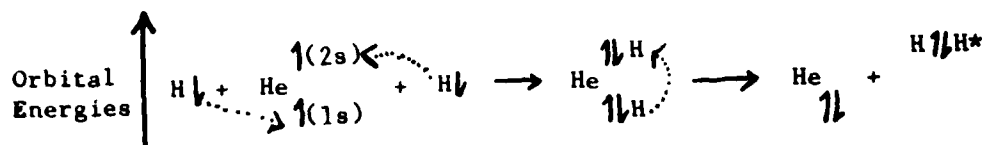


FIGURE 3. PRESUMED DECAY SCHEME OF  $\text{HeH}_2$ .

surfaces when the helium approaches the  $\text{H}_2$  center of mass at an angle of  $45^\circ$ . A similar minimum is recognized for an He ... H-H unsymmetrical configuration. These are excited molecular states, however, so further stabilization, perhaps through a collective mechanism, would be necessary for long-term stability. The third triatomic species represents a system which emphasizes the possibility of hydrogen bond character in the molecule. None of these systems have been experimentally isolated.

Table IV presents several homonuclear helium systems. The first

SYSTEM	ELECTRONIC STRUCTURE	FORM OF WAVE FUNCTION	ORBITAL POPULATION
$\text{He}_2^{++}$	$1\sigma_g^2 ({}^1\Sigma_g^+)$	<p>Energy ↑</p> <p>He atom      Molecule      He atom</p> <p><math>1\sigma_u (AB)</math></p>	— $\uparrow\downarrow$ $\uparrow\downarrow$
$\text{He}_2^+$	$1\sigma_g^2 1\sigma_u ({}^2\Sigma_u^+)$	<p>He atom      Molecule      He atom</p> <p><math>1\sigma_g (B)</math></p>	$\uparrow\downarrow$ $\uparrow\downarrow$ $\uparrow\downarrow$
$\text{He}_2$	$1\sigma_g^2 1\sigma_u^2 ({}^1\Sigma_g^+)$	<p>He atom      Molecule      He atom</p>	$\uparrow\downarrow$ $\uparrow\downarrow$ $\uparrow\downarrow$ $\uparrow\downarrow$
$\text{He}_6$	$(D_{6h})$	<p>(Non-valence)</p>	

TABLE IV. HOMONUCLEAR HELIUM SYSTEMS. See Table III for explanation of symbols.

three species are well-known from the earliest studies in molecular quantum mechanics. Like the diatomic hydride systems, only the singly charged system is energetically stable, the first one spontaneously dissociating into two  $\text{He}^+$  ions and the third one possessing a slightly

repulsive potential in the ground state. The last system,  $\text{He}_6$ , represents an intriguing prediction of 25 years ago (16) for van der Waals stability. To this author's knowledge, this system has never been experimentally identified.

Table V predicts possible halide systems with helium. Note that p-orbitals are here invoked for the first time. The first two systems indicate the form in which s and p orbitals combine and also represent the simplest theoretical test of the strength of the helium ionization potential against the electronegativity of elemental fluorine. As yet,  $\text{HeF}^+$  and  $\text{HeF}$  are not known experimentally.  $\text{He}_2\text{F}^+$  is of interest because of the possibility of sharing helium atom electrons with a very electronegative  $\text{F}^+$  ion. A counter ion might then be found to neutralize or stabilize the system. Unfortunately, the  $\text{F}^+$  ion cannot be produced by any chemical reaction.  $\text{HeF}_2$  is included because of its obvious similarity to the extremely stable  $\text{F}-\text{H}-\text{F}^-$  bifluoride ion (-37 kcal/mole). Although the  $\text{Cl}-\text{H}-\text{Cl}^-$  dichloride ion is also stable (-14 kcal/mole), it is not believed that less electronegative halogens will be likely to form stable helium species. The character of bonding with chlorine would be very similar to that with fluorine, in any event. An extensive set of theoretical references (17 - 21) is available on helium-fluorine systems because of the critical electronegativity value for fluorine.

Table VI presents several helium-oxygen systems which we have considered. These systems warrant further investigation for different reasons than the previously listed molecular systems. Not only is oxygen highly electronegative, but here we deal with two different oxygen species, both of which have stable unpaired electron configurations, hence an electron pairing with helium electrons intuitively seems possible. There has been some speculation (22) that the three oxygen species is also a diradical, but an additional interest is because of the well-known oxidizing power of ozone. Although a naive bonding picture is not readily apparent, a more sophisticated bonding description has been attempted. A preliminary MO calculation was made, based on symmetry orbitals arising from the three oxygen atoms

SYSTEM	FORM OF WAVE FUNCTION			ORBITAL POPULATIONS	
$\text{HeF}^+$ ( $C_{\infty v}$ )				—	+
$\text{HeF}$ ( $C_{\infty v}$ )					
	He atom	Molecule	F atom	$\text{HeF}^+$	$\text{HeF}$
$\text{He}_2\text{F}^+$ ( $D_{\infty h}$ )				—	
	He atom	F atom	He atom		
$\text{HeF}_2$ ( $D_{\infty h}$ )				—	
	F atom	He atom	F atom		

TABLE V. HELIUM HALIDE SYSTEMS. See Table III for explanation of symbols.

and helium in a  $C_{2v}$  point group symmetry. Approximations were chosen for Coulomb and resonance integrals and the secular equation was solved for nine roots. Unfortunately, two roots proved unrealistically complex,

## SYSTEM

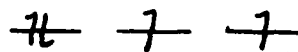
## FORM OF WAVE FUNCTION AND COMMENTS

HeO  
(C<sub>∞v</sub>)

$$\psi = c_1 1s_{\text{He}} + c_2 (2p_{x,0} + 2p_{y,0})$$

$$\psi = c_1 \left( \begin{array}{c} + \\ \circ \end{array} \right) + c_2 \left( \begin{array}{c} + \quad - \\ \circ \quad \circ \end{array} \right) + \left( \begin{array}{c} + \\ \circ \\ - \end{array} \right)$$

Note that oxygen atom has an unpaired 2p electron population:



HeO<sub>2</sub>  
(C<sub>2v</sub>)

(bent)

Oxygen molecule also has unpaired electrons in two degenerate, singly occupied MO's:



HeO<sub>3</sub>  
(C<sub>2v</sub>)

(bent)

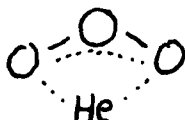


TABLE VI. HELIUM OXYGEN SYSTEMS.

presumably because of the infelicitous choice of empirical values for the integrals. This calculation is being repeated, with a more careful selection of integral values.

Some theoretical calculations indicate stable states of alkali-metals with helium, as molecules and ions. An especially extensive set of calculations on LiHe have been reported (23).

A set of neon systems has been considered, also, primarily because of the markedly lower ionization potential of neon (21.6 eV) compared to helium (24.6 eV) and the consequent improved possibility of binding in NeH<sub>2</sub>; NeO, NeO<sub>2</sub>, NeO<sub>3</sub>; and NeF<sub>2</sub>.

Finally, we mention additional systems that might form, either through an attraction to exposed protons (by virtue of a high electro-negativity counter atom) or through a coordinate covalence. These systems include He ... H<sub>3</sub>N, He ... H<sub>3</sub>N<sup>+</sup>, He ... H<sub>2</sub>N, He ... BH<sub>3</sub> and He ... BF<sub>3</sub>.

#### IV. CONSIDERATIONS OF SYNTHETIC PATHWAYS

Numerous procedures have been proposed for synthesizing rare gas systems. We have organized several categories of synthetic approaches,

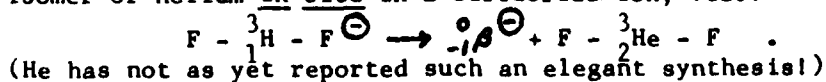
specifically, chemical, molecular beam and radiation excitation.\*

The first type of chemical synthesis considers the possibility of direct oxidation of helium by a suitably strong oxidizing agent which can then furnish a combining counter-ion. Literature references hint at the following as likely candidates(14): F, F<sub>2</sub>; O, O<sub>2</sub>, O<sub>3</sub>; O<sub>2</sub>F<sub>2</sub>, Cl<sub>2</sub> and N<sub>2</sub>. All of these oxidizers have large electronegativity values and several have been proposed in the recent literature (24). Positive synthetic results have not as yet been reported, however. A second, less popularized chemical approach would use a Lewis acid to attempt a coordinate complexing, as with the He ... BF<sub>3</sub> proposal above. Success in this area has also not yet been reported. Although exhaustive experiments have not been undertaken with all oxidants or Lewis acids, serious reasoning concerning ionization potentials and electronegativities does not give strong encouragement to these approaches, at least from the viewpoint of theoretical arguments, without the possibility of an additional stabilization, perhaps through some cooperative physical effects. We still believe that some selected experimentation along chemical lines is warranted.

Molecular beams (25) and adiabatic nozzle expansions (26) have been proposed as means of forming particular, stable states and/or furnishing additional stability by rapidly dropping the system into a very low vibrational level of the electronic state. These experiments undoubtedly are of a more sophisticated nature than the prospect of a relatively simple chemical reaction. These procedures produce various transient species, so additional experiments should be designed, for instance He\* + O, He\* + O<sub>2</sub>, He\* + O<sub>3</sub>, where He\* represents independently excited helium atoms, in order to seek particularly attractive prospective molecular systems.

Radiation excitation can be of two types, saturation with high

\*We exclude certain exotic proposals. Pimentel has suggested the intriguing possibility of a  $\beta$ -decay of a tritium atom to form the mass-3 isomer of helium in situ in a bifluoride ion, viz.:





intensity microwave or radiofrequency fields (27) to provide a pool of excited atoms and molecules for collisional association and "dressing" (28) of atoms or molecules with high intensity lasers, usually conceived as a means of shifting certain potential energy surfaces with respect to others. Laser "dressing" has the effect of lowering threshold energies for electronic transitions and even causing potential curves to cross where normally they are separated by energies of the order of an electron volt. These methods are more speculative than chemical or beam methods, so require more study before their utility can be properly assessed.

An additional scheme of preparing excited reactants, within a beam experiment or in a flowing gas reactor, is the technique of x-ray excitation which has proven itself in several laboratories as an efficient means of exciting systems into metastable levels (29).

#### V. RECOMMENDATIONS

The preliminary work reported here should be continued in several directions. First, sophisticated computer calculations should be initiated on some of the more promising molecular systems in order to determine whether stable configurations exist. Initially, ground state structures should be investigated for potential surface minima but eventually excited state configuration calculations should be warranted. (GAUSSIAN 76 is a recommended ab initio computational program although HONDO 5 is a more elaborate possibility -- both are available from the Quantum Chemistry Program Exchange, Indiana University, in forms compatible with the CDC CYBER at the Rocket Propulsion Laboratory.) Likely candidates for calculation include  $\text{HeO}^+$ ,  $\text{HeO}$ ,  $\text{HeO}_2$  and  $\text{HeO}_3$ . Estimates of approximately 10 minutes CPU computing time, per geometrical configuration, for  $\text{HeO}_2$  have been given to us by QCPE spokesmen. A second direction for continued research is a spectroscopic search for stable states of high energy systems, which is complementary to the computer studies. Relatively simple experiments using lasers, flash lamps or radiofrequency excitation of chemical components can be followed by moderate dispersion spectral investigation to seek evidence for the existence of high energy species. Again, the helium fluoride

and helium oxide systems are likely prospects, as are the helium-ammonia and helium-boron trifluoride mixtures. An inexpensive, low-resolution mass spectrometer may simultaneously be employed for gathering confirmatory evidence of a system's stability.

The actual work suggested here is relatively straightforward. The analysis of data, either to search for more likely configurations for theoretical computations, or to detect unique spectroscopic signatures of novel species, will require sophisticated background and insight. It may prove possible to fit calculated spectra with observed features, if structural parameters are effectively chosen. Indeed, spectroscopists have been able to accurately calculate rotational-vibrational envelopes of electronic transitions for some time (30). Such calculations also require large, main-frame computers such as the Rocket Propulsion Lab possesses.

Eventually, molecular- and ion-beam machines may be desired to effect synthetic collisions of components into preferred electronic states. If the first two parts of the program (calculations and diagnostic spectroscopy) have shown promise, then an additional order-of-magnitude funding should be warranted\*. A recent paper (31) effectively illustrates how the three approaches of spectroscopy, beam experiments and ab initio calculations can complement each other in assigning and characterizing attractive potential wells, in this case for excited, stable states of  $\text{He}_2$ .

---

\*It has been gratifying to discuss some of these options with Dr. Ronald Herm, an experienced molecular beam researcher, who appears to be willing to consult on the fabrication of an appropriate beam machine and the design of critical experiments along the lines which we have discussed above.

## REFERENCES

1. E. Wigner and H.B. Huntington, "On the Possibility of a Metallic Modification of Hydrogen", J. Chem. Phys. Vol. 3, 764 - 770 (1935).
2. J. S. Zmuidzinas, "Spin-Polarized Triplet Helium", unpublished Technical Report from U.S. Air Force Rocket Propulsion Lab, AFRPL TR-84-027, April, 1984, 41 pages.
3. R.F. Ferrante and W.C. Stwalley, "Spin-Polarized Atomic Nitrogen and the  $\sum_u^+$  State of  $N_2$ ", J. Chem. Phys. Vol. 76, 3107 - 3111 (1983).
4. J. de Boer, "Quantum Theory of Condensed Permanent Gases. I. The Law of Corresponding States.", Physica, Vol. 14, 139 - 148 (1948).
5. J.R. Woodworth, "Determination of the Spontaneous Radiative Transition Rate of the  $2^1S \rightarrow 1^1S$  Transition in Neutral Helium", Thesis, Johns Hopkins Univ., 166 pages, [Dissertation Abstracts Int. 35B, 6024 (1975)].
6. P.W. Atkins, Molecular Quantum Mechanics, 2<sup>nd</sup> Ed'n (Oxford Univ. Press, New York, 1983), Chapt. 10.
7. S.P. McGlynn, L.G. Vanquickenborne, M. Kinoshita and D.G. Carroll, Introduction to Applied Quantum Chemistry, (Holt, Rinehart and Winston, Inc. New York, 1972).
8. P. Jørgensen and J. Simons, Second Quantization-Based Methods in Quantum Chemistry, (Academic Press, New York, 1981).
9. R.S. Mulliken and W.S. Ermler, Diatomic Molecules, Results of Ab-Initio Calculations, (Academic Press, New York, 1977).
10. K.P. Huber and G. Herzberg, Molecular Spectra and Molecular Structure, Vol. IV, Constants of Diatomic Molecules, (Van Nostrand-Reinhold Co., New York, 1979).
11. R.S. Mulliken and W.S. Ermler, Polyatomic Molecules, Results of Ab-Initio Calculations, (Academic Press, New York, 1981).
12. G. Herzberg, Molecular Spectra and Molecular Structure, Vol. III, Electronic Spectra and Electronic Structure of Polyatomic Molecules, (Van Nostrand-Reinhold Co., New York, 1966).
13. See the discussion in J.C. Slater, Quantum Theory of Atomic Structure, Vol. II, (McGraw-Hill Book Company, New York, 1960), especially Chapter 18.
14. An excellent bibliography of early experimental and theoretical work

is contained in J.H. Holloway, Noble Gas Chemistry, (Methuen, London, 1968), especially Part II. Mulliken and Ermler (9) carry theoretical results up through the middle 1970's and a bibliography to March, 1983 is being compiled for a U.S. Air Force Rocket Propulsion Lab Tech Memo which is a companion to this chapter. Much of this material is taken from the Berkeley Molecular Spectra Newsletter, collated and distributed by J.G. Phillips and S.P. Davis.

15. S.C. Farantos, G. Theodorakopoulos and C.A. Nikolaides, "A Non-van der Waals Minimum of the  $\text{He}(^1\text{S}) + \text{H}_2(\text{B}^1\Sigma^+)$  Excited Surface", Chem. Phys. Letters Vol. 100, 263 - 264 (1983).
16. N. Bernardes and H. Primakoff, "Molecule Formation in the Inert Gases", J. Chem. Phys. Vol. 30, 691 - 694.
17. G.C. Pimentel, "The Bonding of Trihalide and Bifluoride Ions by the Molecular Orbital Method", J. Chem. Phys. Vol. 19, 446 - 448 (1951).
18. G.C. Pimentel and R.D. Spratley, "The Bonding in the Inert Gas-Halogen Compounds -- The Likely Existence of Helium Difluoride", J. Am. Chem. Soc. Vol. 85, 826 - 827.
19. L.C. Allen, R.M. Erdahl, and J.L. Whitten, "Helium Difluoride", J. Am. Chem. Soc. Vol. 87, 3769 - 3771 (1965).
20. L.C. Allen, A.M. Lesk and R.M. Erdahl, "The Fluorides and Oxides of Helium and Neon", J. Am. Chem. Soc. Vol. 88, 615 - 616 (1966).
21. R.K. Nesbet, "Theory of Inert-Gas Fluorides", J. Chem. Phys. Vol. 38, 1783 - 1784 (1963).
22. R.J. Hay, T.H. Dunning, Jr. and W.A. Goddard, III, "Configuration Interaction Studies of  $\text{O}_3$  and  $\text{O}_3^+$ . Ground and Excited States", J. Chem. Phys. Vol. 62, 3912 - 3924 (1975).
23. S.B. Schneiderman and H.H. Michels, "Quantum Mechanical Studies of Several Helium-Lithium Interaction Potentials", J. Chem. Phys. Vol. 42, 3706 - 3719 (1965).
24. J.F. Liebman and L.C. Allen, "A Salt Chemistry of Light Noble Gas Compounds", J. Am. Chem. Soc. Vol. 92, 3539 - 3543 (1970).
25. D.H. Stedman and D.W. Setser, "Chemical Applications of Metastable Rare Gas Atoms", Prog. in Reaction Kinetics, Vol. 6 (Pergamon, New York, 1972), p. 193 - 238.
26. R.E. Smalley, L. Wharton and D.H. Levy, "Molecular Optical Spectroscopy with Supersonic Beams and Jets", Accts. Chem. Res. Vol. 10, 139 - 145 (1977).

27. See G.W. Robinson, "Electronic Spectra", Chapt. 2.4 in D. Williams, Ed., Methods of Experimental Physics, Vol. 3, Molecular Physics (Academic Press, New York, 1962).
28. See, for example, A. Ben-Shaul, Y. Haas, K.L. Kompa and R.D. Levine, Lasers and Chemical Change (Springer-Verlag, New York, 1981).
29. S.A. Goldstein, A.P. D'Silva and V.A. Fassel, "X-ray Excited Optical Fluorescence of Gaseous Atmospheric Pollutants: Analytical Feasibility Study", Radiation Res. Vol. 59, 422 - 437 (1974).
30. J.H. Callomon, T.M. Dunn and I.M. Mills, "Rotational Analysis of the 2600Å Absorption System of Benzene", Phil. Transactions Roy. Soc. (London) Vol. A259, 499 - 532 (1966).
31. R.M. Jordan and P.E. Siska, "Potential Energy Curves for the  $A^1\Sigma_u^+$  and  $C^1\Sigma_g^+$  States of  $He_2$  Obtained by Combining Scattering, Spectroscopy and Ab Initio Theory", J. Chem. Phys. 80, 5027 - 5035 (1984).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

Rewrite of AFM 28-345

Prepared by:	Evelyn J. Thompson Leggette
Academic Rank:	Associate Professor
Department and University:	Reading Jackson State University
Research Location:	Logistics Management Center Logistics Plans Gunter Air Force Station, Alabama 36114
USAF Research:	Capt Gail A. Christmas
Date:	September 84
Contract No.:	F49620-82-C-0035

REWRITE OF AFM 28-345

by

Evelyn J. Thompson Leggette

ABSTRACT

The objective of the project was to rewrite selected nontechnical sections of AFM 28-345, Contingency Operation/Mobility Planning and Execution System (COMPES-LOGMOD-B) to make it more readable and easy to understand.

A thorough review of the literature was conducted on readability; Fry's Readability Graph was applied to selected sections to ascertain the appropriate grade level of the manual; a questionnaire was developed to determine the opinions of the users of the manual relative to format, sentence construction, vocabulary, reading and educational levels; and, any other suggestions for improvement.

Based on the review of the literature, the ascertained grade level and the results of the questionnaire, revisions have been made for the manual, and a grade level has been determined for the revisions. Recommendations are also offered for improvement and a follow-on study on other aspects of the manual has been suggested.

#### ACKNOWLEDGEMENT

The writer would like to express her appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to spend a very challenging, informative, and rewarding summer with the Logistics Plans Directorate of the Air Force Logistics Management Center, Gunter Air Force Station, Alabama.

Special appreciation is offered to Captain Gail Christmas and Lt Col Cecil Smith for their guidance and collaboration. Sincere thanks to Maj Holland, Maj Snyder, Maj Stewart, Capt Grandalski and CMSgt Hankins for their helpful cooperation during this ten week period.

Acknowledgement is also accorded Ms Kay Booth for the typing of this research report.



## Rewrite of AFM 28-345

### I. INTRODUCTION

The ability to grasp and understand printed information with minimum difficulty is vital to job effectiveness and success. This ability is very fundamental and is an extremely important concern to persons who must consume and apply volumes of printed information for job related tasks. On 30 Sep 1975, the Air Force Logistics Management Center (AFLMC) was activated at Gunter Air Force Station, Alabama and assigned to Headquarters Air University, Maxwell Air Force Base. Its mission was to develop logistics concepts and procedures for improving Air Force logistics system. The Directorate of Logistics Plans is responsible for Air Staff-directed studies, evaluations, and analyses of existing and proposed functional management processes. This involves evaluating current logistics planning management systems to identify areas which require improvement or enhancement to increase efficiency and effectiveness of personnel, materiel, procedures, and process resources.<sup>1</sup>

One of the projects of Logistics Plans is the development of AFM 28-345 War Planning-Contingency Operation/Mobility Planning and Execution System (COMPES) Logistics Module - Base-Level (LOGMOD-B) Users Manual. According to AF Manual 28-345 "the manual provides the guidance necessary for Air Force logistics planners to automate base-level mobility planning in support of contingency operations. The system provides standard procedures governing the preparation of input data to the computer. This manual establishes system/program responsibilities,

procedures for establishing the COMPES LOGMOD-B programs and the data records required for maintaining accurate mobility information. This manual applies to those Air Force units, active and reserve forces, that deploy, receive or support Air Force forces with a mobility commitment."<sup>2</sup> The objective of the Users Manual for COMPES is to provide mobility planners with the information necessary to effectively use the automated system at unit/base level. Because of the role that logistics plans play in providing mobility planners with information to use effectively the automated system at unit/base level, a rewrite of AFM 28-345 was initiated.

The present project is concerned with rewriting selected nontechnical sections of AF Manual 28-345 users manual for increased clarity and understanding. To begin the project, the manual was read thoroughly to determine the difficulty of the information. Special attention was given to the length of the sentences, the arrangement and order of the sentences, the multisyllabic words, and the number of long or unusual words in the manual. These factors were considered mainly because they sometimes inhibit comprehension and understanding of printed information. Additionally they are factors used in determining readability levels of printed information.

Two readability formulas applied to random sections of the manual ascertained an approximate reading level of 14 which suggests that the printed information is geared toward a college level audience. The Fry's Readability Graph was applied to determine the approximate difficulty level of the User's Manual. However, there are several readability formulas that can be used to determine the approximate grade

level of written materials, or more accurately, to compare the relatively difficulty of materials. Fry's Readability Graph is designed to measure printed material from the fifth grade through graduate level<sup>3</sup>; thus, it is deemed appropriate for determining the approximate difficulty level of AF Manual 28-345.

To determine the opinions of the users of AFM 28-345, and substantiate revisions incorporated in the manual by the researcher, a questionnaire was developed and sent to selected users and MAJCOM representatives relative to format, sentence construction, vocabulary, reading and educational levels of the manual. Hence, the actual nontechnical revisions and rewrite of the manual were based on the experiences and training of the researcher, the review of the literature, and the results (where possible) of the data generated from the questionnaire.

In Section II the objectives of the research are stated. Sections III, IV, V present the review of the literature, approaches taken in achieving each of the objectives, and the results of each approach. Section VI offers recommendations for implementing results of the research and follow-on research considerations.

## II. OBJECTIVES OF THE RESEARCH EFFORT

The objective of this project was to rewrite selected nontechnical sections of AFM 28-345, Contingency Operation/Mobility Planning and Execution System (COMPES LOGMOD-B) to make it more readable and easy to understand. The specific objectives to be explored are:

- 1) To review the literature on readability.
- 2) To determine the present reading level of selected sections

of AFM 28-345 (COMPES LOGMOD-B).

- 3) To develop a questionnaire to determine specific problems on instructions and understanding of AFM 28-345.

In order to accomplish these objectives, the following approaches were undertaken. (1) A thorough search of the literature was conducted on readability and revision of printed information. Pertinent research will be reported in Section III with emphasis on a rationale for selected Fry's Readability Graph to utilize in this research investigation.

- 2) The researcher examined AFM 28-345 by reading and studying the manual to determine: the reading difficulty level from a layman's perspective; the logical organization and arrangement; the length and construction of sentences; and, the vocabulary load. After several readings of the manual, Fry's Readability Graph was applied to random sections of it to ascertain an approximate grade level. The readability formula and its applications are discussed in Section IV.

- 3) A questionnaire (See Appendix A) was developed to gain input from users of AFM 28-345. The first draft was submitted to the staff here in Logistics Plans to scrutinize it for a) if the instructions were clear; b) if the questions were clear and concise; c) if the questions were pertinent; and, d) what additional information was needed to make it more complete. Based on the comments and suggestions for the improvement of the instrument, the final draft was completed. An analysis of the questionnaire will be detailed in Section IV-A. A series of rewrites and revisions will be discussed briefly in Section V. Recommendations are cited in Section VI.

### III. REVIEW OF THE LITERATURE

Readability is the objective measure of the difficulty of a book or article of printed information. Factors that influence readability have been the subject of scholarly study for thousands of years. According to Klare, "Lorge tells of words and idea counts made by Talmudists in 900 A.D. so that they could use frequency of occurrence to distinguish from usual senses (meanings).<sup>4</sup>

Although the study of readability has been underway for many years, little progress was made until statistical techniques were developed. Such techniques made it possible to identify important readability factors and to construct formulas for estimating passage difficulty.

Several dozen readability formulas have been developed over the years. Many of the earlier formulas were complex and required the user to count a number of variables. Later research revealed that the two most important factors influencing readability of commonly printed materials in the United States are sentence length and word difficulty.<sup>4</sup> In general, the longer the sentences and the words, the more difficult is the material to read. Conversely, if the sentences are short and the words are short, the material is easier to read.

Today, the more commonly accepted and widely used formulas use sentence length and word difficulty for estimating reading level. Examples of this method are the Spache Readability Formula for primary grade materials and the Dale-Chall Readability Formula for materials in grades four through sixteen. Of the less time-consuming formulas, Fry's Graph of Estimating Readability is one of the most commonly accepted. His graph is an outgrowth of a need for a simple and efficient technique

for estimating reading difficulty. Traditional factors of sentence length and word difficulty are used to determine readability. Fry<sup>3</sup> reports that the graph correlates 0.90 with Spache Readability Formula and 0.94 with the Dale-Chall Readability Formula. These high correlations indicate considerable consistency between formulas and support the wide acceptance and use of Fry's graph. Fry's procedure involves selecting three or more 100-word passages, counting the number of syllables in the words in those passages, and determining the length of the sentences contained in the passages. The averages of the syllable counts and sentence lengths are plotted on a graph to determine the approximate grade level.

The Fry technique estimates the reading ability one needs in order to understand the material with 50% to 75% accuracy. A Fry grade level score of 10 indicates that a tenth grader could read this material but would need some assistance.<sup>6</sup>

Psycholinguistic factors also affect the difficulty level of materials. Is the reader familiar with the language pattern used by the writer, e.g., phrasing and syntax, semantics, and the paragraph, chapter and book organization? Specific linguistic factors related to syntax which seem important in readability are sentence patterns and intonation clues. Normal subject-to-predicate patterns (Sam arrived late) are probably easier to read than are transformations of these patterns (Late arrived Sam). Semantics, or word meanings also needs consideration. What makes a word easy or difficult to read and understand? Is word length a factor? Are multiple or remote meanings a factor? How are paragraphs organized? What kinds of illustrations, if any, are used?

What kind and size print is used? All of these factors affect the difficulty of printed materials.<sup>7</sup> In most cases, there is a definite interaction between factors found in written materials and factors found within the person who is reading.

Critics of readability procedures feel there is more to estimating the ease of or difficulty of printed information than numbers or syllables per word and sentence length. Collins<sup>8</sup> and his colleagues mention "extra linguistic variables," which include illustrations, reader interest, reader knowledge, and writer's purpose and style. Leggette<sup>9</sup> alludes to changing polysyllabic to non-polysyllabic words by selecting appropriate synonyms that may be substituted without paraphrasing and risking a loss of content.

Revising the sentence structure of the selection is perhaps the easier of two ways to alter the readability level. Often good materials include long sentences with many details used to express the thought or idea. To lower the readability, it is necessary to introduce the facts in many sentences rather than in one long complex sentence. This aids the reader's organization of facts and usually improves comprehension because the facts are introduced in shorter doses. Another method of lowering readability is to decrease the number of syllables by substituting shorter words for longer words. A dictionary or thesaurus is helpful when changing words in the selection; however, it is difficult to find synonyms for some basic words. In this case, the emphasis should be on changing the sentence length rather than on changing vocabulary.

Steps for altering readability to a lower level include:

- Step 1 Look for difficult words in the selection.
- Step 2 Change some of the harder words to easier words.
- Step 3 Reread some of the sentences to determine if you can divide some of them into two or three shorter sentences.
- Step 4 Rewrite the selection using shorter sentences and easier words.
- Step 5 Determine the readability of the altered selection to see if you have lowered its reading level.<sup>6</sup>

#### IV. RESULTS OF EACH APPROACH

##### Fry's Readability Graph

Fry's Readability Graph was applied to selected sections of AFM 28-345 to determine the present reading level. Sample passages from Section 3 are presented in Table 1.

TABLE 1  
SAMPLE  
CALCULATION FROM THE FORMULA

FRY'S TECHNIQUE	NUMBER OF SENTENCES	NUMBER OF SYLLABLES
Passage 1	4.2	188
Passage 2	3	151
Passage 3	3.7	165
Average	3.6	168
Readability	14th Grade Level	

\*Section 3 - COMPES

The following passages from Section 3 of AFM 28-345 revealed that the manual is designed for an audience who has completed two years of



college. Passage 1 utilized 3.2.4.14 TAILOR KEY; Passage 2 utilized 3.2.4.46 UPLOAD UNIT TYPE CODE (UP UTC); and, Passage 3 utilized 3.6 LOGMOD-B BACKUP PROCEDURES.

#### PASSAGE 1

##### 3.2.4.14 Tailor Key (TL Key).

a. Cause of input - As needed, to specify tailor keys associated with item priorities, priority increments, deployment echelons and/or UTCs which are to be tailored out (not printed or summarized) during deployment execution. During preplanning, the logistics functional manager reviews the OPlans Op Ords, etc. As a result of the review, items (type and quantities) within a specific unit type code may be directed to be omitted to satisfy a specified tasking. The tailoring of items from a requirements should be coordinated with the respective work center supervisor and squadron mobility officer/NCO prior to such action. Refer to subparagraph.....

#### PASSAGE 2

##### 3.2.4.46 UPLOAD UNIT TYPE CODE (UP UTC)

c. Origin of input - When a new UTC not belonging to an existing group or a group of UTCs not previously defined in the wing's file is needed, the logistics LOGMOD-B functional manager should refer to the LOGMOD Data Transfer Tape Index, PCNSA200-354, to determine if the standard UTC is available. If the tape index is not on hand, the user should prepare the PR PUD for input to obtain the tape index. After it is determined that the UTC(s) required is available and can be uploaded (meets the criteria), the logistics LOGMOD-B functional manager prepares the UP UTC.....

#### PASSAGE 3

##### 3.6 LOGMOD-B BACKUP PROCEDURES

LOGMOD-B Backup Procedures. There may be times when the supporting computer is not operational, time does not permit an immediate computer update cycle and/or an error is detected in a LOGMOD-B output product. Pen and ink changes to existing output products may be necessary during an exercise or deployment to reflect corrected data. When these changes are permanent, the LOGMOD-B data base should be corrected during the next update. In summary, when problems exist in the LOGMOD-B data base or data is invalid on current output products, the user must be prepared to revert to.....(2)

Sentences and syllables are counted and averaged according to Fry's formula. In Table 1, Passage 1 has 4.2 sentences and 188 syllables. Passage 2 has 3 sentences and 168 syllables. Passage 3 has 3.7 sentences and 168 syllables. These three passages average 3.6 sentences and 168 syllables. Using Fry's Readability Graph to plot the average number of sentences and syllables, a readability of 14th grade level is obtained. (See Appendix B).

Fry's technique estimates the reading ability one needs in order to understand the information with 50% to 75% accuracy. If the users of the manual reading grade level is approximately second year college level, he could use the information but would need some assistance from those experienced with the manual.

In general, however, readability formulas are not meant to pinpoint the exact grade level of passages. Instead, they are designed to show the relative difficulty of passages and/or to give the range of difficulty among passages.

A second readability formula (Gunning-Fog) was applied and yielded an average grade level of 14.4 (See Appendix C).

#### IV-A Questionnaire

After extensive reading and studying the user's manual; an interview with a current user; and, suggestions from the Logistics Plans staff, a questionnaire was developed to determine the problem areas of AFM 28-345 relative to format, instructions, sentence length, vocabulary and educational levels. The questionnaire included an introductory message and eight (8) questions with a request to provide comments for each question.

The Message Distribution Center distributed the questionnaires and received responses from users and MAJCOM representatives of COMPES. Thirty-four (34) users and MAJCOM representatives were randomly selected for inclusion in the project. Out of the thirty-four selected, twenty-eight responded giving an eighty-two (82) percent return.

Table 2 presents an analysis of the opinions of the respondents. Fifty percent of the respondents indicated "yes" and the same percentage "no" for Question A. For the "yes" responses, the general comment was that the manual is clear and logical with some minor exceptions. Those who indicated that the manual was not clear and legible made reference to the small and difficult print, too many sub-paragraphs and too many notes.

He recognizes that there are exceptions to this generalization, but feels that the number of syllables a word contains is usually a good index of its difficulty.<sup>11</sup>

Most respondents agree that too many acronyms are in the manual and one has to refer back for definitions of acronyms while using the manual; the relationship between various acronyms are not explained; acronyms slow reading and overall comprehension; and they present a problem to new users. Almost fifty percent of the respondents agree that they limit understanding and memory.

Seventy-five (75) percent feel the reading level is not too difficult for the users of the manual. The biggest problem is that the manual is written by computer technicians rather than logisticians. However, it is felt that the manual does meet Air Force Standards relative to the reading level. As can be seen in the table, the educational level appropriate for users of the manual is twelfth grade.

Yet, the person who has completed twelfth grade does not necessarily read and comprehend at a twelfth grade level, especially printed information of a technical, non-visual, and non-graphic nature. This in itself suggests a need for adjusting the readability to a suitable level.

Other additional comments by the respondents include a) AFM 28-345 is a well-written manual; it provides good guidance on the technical operation, but not on data base structure and system management. b) More crosstells like the questionnaire and formal training sessions for the logisticians could alleviate the learning experiences that come through trial and error. c) The manual is too large and should be

The legibility and visibility of printed materials influence its readability. Factors such as line length, style of type face, space between lines and between letters, margins and physical format are important to legibility; whereas, color, finish and print color are important to visibility. These "variables" allow the sensory process to operate efficiently, and thus form a basis for the perceptual process of comprehensibility or readability.<sup>10</sup> Less than fifty percent of the respondents feel the instructions are easy to understand and fifty-seven (57) percent indicated that they are not easy to understand.

The percentage for responding to Question C is fifty percent (50) for "yes" and "no." Some comments include most sentences seem to be of a length that can be understood; other comments indicate that long sentences should be changed to shorter sentences.

Linguistics factors related to syntax and are important in readability are sentence patterns and intonation clues. Regular subject-to-verb patterns are easier to read than are transformation of these patterns, particular when knowledge and understanding is to be gained from the information.

Relative to Question D, "Does the vocabulary present a problem in understanding?", the majority of the respondents indicated "yes." It is felt by these respondents that a large technical vocabulary and computer background are necessary to understand the terms. Those responding "no" said any new system undergoes a learning process and one should utilize the abbreviation section of the manual to help understand the terms.

According to Flesch, a word is difficult or easy according to the number of syllables it contains: the more syllables, the harder it is.

He recognizes that there are exceptions to this generalization, but feels that the number of syllables a word contains is usually a good index of its difficulty.<sup>11</sup>

Most respondents agree that too many acronyms are in the manual and one has to refer back for definitions of acronyms while using the manual; the relationship between various acronyms are not explained; acronyms slow reading and overall comprehension; and they present a problem to new users. Almost fifty percent of the respondents agree that they limit understanding and memory.

Seventy-five (75) percent feel the reading level is not too difficult for the users of the manual. The biggest problem is that the manual is written by computer technicians rather than logisticians. However, it is felt that the manual does meet Air Force Standards relative to the reading level. As can be seen in the table, the educational level appropriate for users of the manual is twelfth grade.

Yet, the person who has completed twelfth grade does not necessarily read and comprehend at a twelfth grade level, especially printed information of a technical, non-visual, and non-graphic nature. This in itself suggests a need for adjusting the readability to a suitable level.

Other additional comments by the respondents include a) AFM 28-345 is a well-written manual; it provides good guidance on the technical operation, but not on data base structure and system management. b) More crosstells like the questionnaire and formal training sessions for the logisticians could alleviate the learning experiences that come through trial and error. c) The manual is too large and should be

separated into parts and chapters - this will enhance comprehension; d) the manual should be written in the language of the laymen, rather than in computer lingo by a computer programmer; e) a section on data system management should be added; f) the manual does not provide a sequential process for using the base level LOGMOD-B system; g) there are too many cross references to enable the user to clearly recall information from one reference as it relates to another, and; h) instructions should contain a building block format of sequential instructions for accomplishing each task in LOGMOD-B. For example, page 1-11 through 1-19 should be in Section 3 with the procedural instructions on priority increments.

#### V. Rewrite and Revisions

There are several steps that are recommended to be followed in rewriting printed information. These steps along with the researcher's experiences are incorporated for the revisions in AFM 28-345.

STEP 1: Look for difficult words in the selection.

STEP 2: Change some of the "harder words to easier words."

STEP 3: Reread the sentences to determine if you can divide some of them into two or three shorter sentences.

STEP 4: Rewrite the selection using shorter sentences and easier words.

STEP 5: Determine the readability of the altered selection to see if you have lowered its reading level.<sup>6</sup>

It is not possible to change every difficult word in a selection but the selection can be made easier to read by using synonyms for some of

the more difficult words. Sentences that are very lengthy with introductory clauses and long phrases should be divided into shorter sentences. They should also be written in the natural order. It is also worthy to note that writing for understanding should be done in much the same way as talking to someone. The main concern is that the ideas are understood.

Table 3 presents samples of rewrites and revisions for Section 3 of AFM 28-345 based on the five-steps for rewriting printed information. Section 3 is utilized because it is the largest section of the manual. For detailed information on rewrites, see revised manual.



Table 3

REWRITES IN SELECTED SECTIONS OF AFM 28-345

Section	Original	Changes (Revisions)
3.2.1.1		
	b. The installation mobility officer, in coordination with squadron, shop and tenant mobility personnel, determines the wing structure for the LOGMOD-B data base.	The installation mobility officer determines the wing structure for the LOGMOD-B data base. This is done with squadron, and shop and tenant mobility personnel.
3.2.1.2		
	LOGMOD-B will process cards/images within a wing's data card/image deck based on the sequence number, not as the cards/images are read through the card/image reader.....	LOGMOD-B will process cards/images within a wing's data card/image deck based on the sequence number. They will not be processed as the cards/images are read through the card/image reader.
3.2.1.2		
	h. If no continuation card/images is needed, position 80 may be left blank. When the "X" is entered, a continuation card/image must follow.	Leave position 80 blank if no continuation card/image is needed. A continuation card/image must follow when the "X" is entered.
3.2.4.46		
	c. When a new UTC not belonging to an existing group or a group of UTCs not previously defined in the wing's file is needed, the logistics LOGMOD-B functional manager should refer to the LOGMOD Data Transfer Tape Index.....	The logistics LOGMOD-B functional manager should refer to the LOGMOD Data Transfer Tape Index, PCN SA 200-345, to determine if the standard UTC is available. This is to be done when a new UTC not belonging to an existing group or a group of UTC(s) not previously defined in the wing's file is needed.

If the tape desk is not on hand, the user should prepare the PR PUP for input to obtain the tape index.

888

After it is determined that the UTC(s) required is available and can be uploaded (meets the criteria), the logistics LOGMOD-B functional manager prepares the UP UTC transaction.

888

If the UTC or UTC(s) are not available on the transfer tape, the logistics function should request the data to be forwarded (if nonpilot unit)<sup>2</sup>

The user should prepare the PR PUP for input to obtain the tape index if it is not on hand.

888

The logistics LOGMOD-B functional manager prepares the UP UTC transaction. This is done after it is known that the UTC(s) is available and can be uploaded (meets the criteria).

888

The logistics function should request the data to be forwarded (if nonpilot unit) if the UTC or UTC(s) are not available on the transfer tape.

---

As can be seen, STEPS 3 and 4 have been utilized in the samples in Table 3. Each sentence was reread to determine if it could be divided into two or three shorter sentences. For 3.2.4.4.6, the sample has at least thirty-nine words. The number of words have been reduced for clarity of expression. In other cases, shorter sentences have been written to control the number of thoughts introduced at one time.

An application of Fry's Readability Graph indicates that the altered revisions yield a grade level of 9 because they now contain 4.0 sentences and 144 syllables. When altering the readability of selections, it is not essential that the work count be exactly the same. Often the word count is increased or decreased when changing vocabulary and sentence length. The goal is to alter the selection so that it may be read and understood.

## VI. RECOMMENDATIONS

Based on the results of this study and the time allocation for completing the objective of rewriting selected nontechnical sections of AFM 28-345, the following recommendations are offered:

1. That on-going research be continued to develop the format, organization and arrangement of the content of the manual since these variables correlate with retention and understanding of information; and, since the manual is constantly being altered.
2. That a logistician and computer programmer be teamed in an on-going research effort to improve other aspects of the manual using the steps and research discussed in the text of the report.
3. That a thorough study on the effects of acronyms on comprehension, memory, and directions be considered since they are used widely throughout the manual.
4. That surveys be conducted with different levels (beginners, current and former) of users of the manual on the revisions. This data will provide a wealth of information to be incorporated in other revisions of the manual.
5. That other variables such as illustrations and visual aids be studied relative to increased clarity and understanding of the manual.
6. That a follow-up study be conducted with the questionnaire respondents to ascertain their opinions of the revisions for AFM 28-345.

APPENDIX A

## QUESTIONNAIRE

1. The Air Force Logistics Management Center (AFLMC) is rewriting selected nontechnical sections of AFM 28-345, Contingency Operation/Mobility Planning and Execution System (COMPES LOGMOD-B) User's Manual, to make it more readable and easier to understand. Working on this project with the LMC is a Summer Research Fellow, Dr. Evelyn J. Leggette. Her expertise is in the field of reading and comprehension skills. To more fully understand the problem areas of AFM 28-345, Dr. Leggette has compiled a questionnaire for all functional users of COMPES LOGMOD-B. Please answer the questions below and provide comments for all "yes" or "no" answers.

- A. Is the format of the manual clear and legible?
- B. Are the instructions arranged in such a manner that they are easy to understand?
- C. Are the sentences too long and wordy for complete understanding?
- D. Does the vocabulary (terms) present a problem in understanding?
- E. Are there too many acronyms in the manual?
- F. Do the acronyms limit understanding and memory of instructions as the manual is being read?
- G. Is the reading level too difficult for the users of the manual?
- H. What education level would probably be appropriate for the users of the manual?

2. Request you provide your response by 17 July 1984. POC is Capt Gail Christmas, AV 446-3535/3355.

APPENDIX B

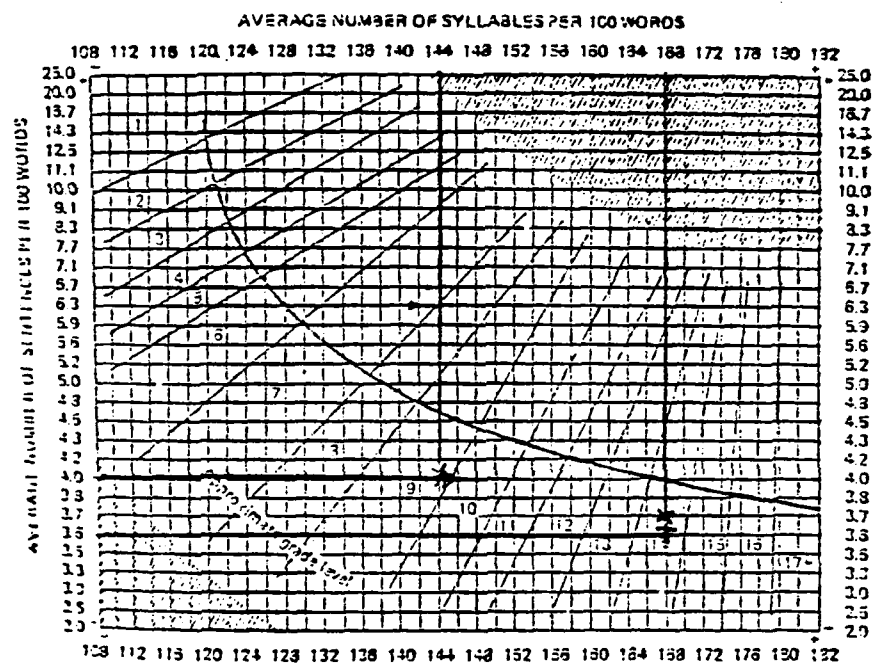


Fig. 22 Edward Fry, "Graph for Estimating Readability—Extended"

Edward Fry, "Fry's Readability Graph: Clarifications, Validity, and Extension to Level 17,"  
*Journal of Reading*, 21, No. 3 (December 1977), p. 249, used by permission of the author  
 and the International Reading Association.

xx before revisions  
 x after revisions

### Expanded Directions for Working Readability Graph

1. Randomly select three (3) sample passages and count out exactly 100 words each, beginning with the beginning of a sentence. Do count proper nouns, initializations, and numerals.
2. Count the number of sentences in the hundred words, estimating length of the fraction of the last sentence to the nearest one-tenth.
3. Count the total number of syllables in the 100-word passage. If you don't have a hand counter available, an easy way is to simply put a mark above every syllable over one in each word, then when you get to the end of the passage, count the number of marks and add 100. Small calculators can also be used as counters by pushing numeral 1, then push the sign for each word or syllable when counting.
4. Enter graph with average sentence length and average number of syllables; plot dot where the two lines intersect. Area where dot is plotted will give you the approximate grade level.
5. If a great deal of variability is found in syllable count or sentence count, putting more samples into the average is desirable.
6. A word is defined as a group of symbols with a space on either side; thus, Joe, IRA, 1945, and & are each one word.
7. A syllable is defined as a phonetic syllable. Generally, there are as many syllables as vowel sounds. For example, stopped is one syllable and wanted is two syllables. When counting syllables for numerals and initializations, count one syllable for each symbol. For example, 1945 is four syllables, IRA is three syllables, and & is one syllable.

---

From Edward Fry. "Fry's Readability Graph: Clarifications, Validity, and Extension to Level 17." *Journal of Reading* 21 (1977):249.



APPENDIX C

## A MODIFICATION OF THE GUNNING FOG INDEX OF READABILITY

### STEP I

Take a sample of approximately 100 words. Get as close to 100 words as possible, but stop at the end of a sentence. Divide the number of sentences into the number of words. \_\_\_\_\_

### STEP II

Count the number of three or more syllable words in the first 100 words of the sample passage with the following exceptions:

- (a) Do not count proper names.
  - (b) Combinations of short easy words like "bookkeeper."
  - (c) Any verb forms, words that are made three syllables by adding "ed" or "es" - like "created" or "trespasses."
- \_\_\_\_\_

### STEP III

Add the two factors above and then multiply by the constant .4, which will give you a grade level. \_\_\_\_\_

### STEP IV

Subtract 1 from the final level.<sup>12</sup> \_\_\_\_\_

NOTE: When counting the three or more syllable words, count the words each time they appear, even if they appear several times in the selection.

AD-A154 337

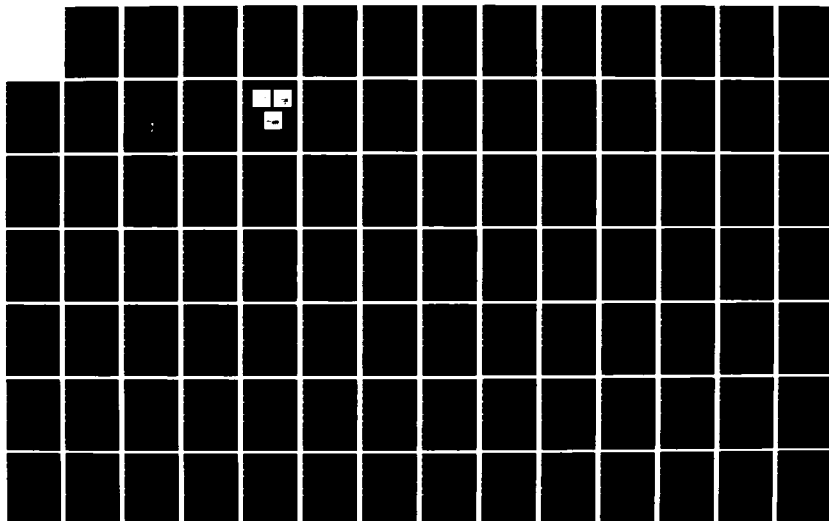
UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 11/13  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

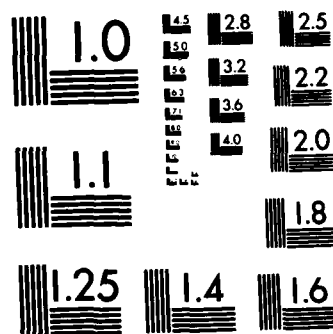
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

## REFERENCES

1. Geiser, Anthony F., Major, USAF Historical Division History of the Air Force Logistics Management Center, January 12, 1976.
2. United States Air Force Manual 28-345.
3. Fry, Edward B., "Fry's Readability Graph: Clarifications, Validity, and Extension to Level 17" Journal of Reading Vol 21, No. 3, December, 1977.
4. Klare, G. R., "The Role of Word Frequency in Readability" Elementary English, Vol 45 (1968) pp. 12-22.
5. Forgan, Henry W. and Mangrum II, Charles, Teaching Content Area Reading Skills 2nd ed Columbus: Charles E. Merrill, 1981.
6. Santensanio, Richard P. A Practical Approach to Content Area Reading Menlo Park, CA: Addison-Wesley Publishing Co., 1983.
7. Burmeister, Lou E. Reading Strategies for Secondary School Teachers Menlo Park, CA: Addison-Wesley Publishing Co., 1974.
8. Collins, Allan M., Ann L. Brown, Jerry L. Morgan and William F. Brewer. The Analysis of Reading Tasks and Texts Technical Report No. 43. Urbana: Center for the Study of Reading, University of Illinois, 1977.
9. Leggette, Earl C. (ed) "SMOG/Leggette Readability Adjuster" (Delta Epsilonian) Phi Delta Kappa Newsletter, 1983.
10. Dechant, Emerald V. and Henry P. Smith, Psychology in Teaching Reading 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1977.
11. Flesch, Rudolf. How to Test Readability. New York: Harper and Row, 1951.
12. Gunning, Robert. The Technique of Clear Writing New York: McGraw-Hill Book Company, Inc., 1952.

1984 USAF-SCEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

NUMERICAL SIMULATION OF A SUPERSONIC INLET FLOW

Prepared by:	Dr. Meng-Sing Liou
Academic Rank:	Visiting Associate Professor
Department and University:	Department of Aerospace Engineering The University of Michigan
Research Location:	Air Force Flight Dynamics Laboratory
USAF Research:	Dr. Wilbur L. Hankey
Date:	September 20, 1984
Contract No:	F49620-82-C-0035

## NUMERICAL SIMULATION OF A SUPERSONIC INLET FLOW

by

Meng-Sing Liou

### ABSTRACT

The progress of computational fluid dynamics has been advanced to a great extent for the past two decades. Various numerical methods have been devised and applied successfully to many complex flows involving both separation and unsteadiness. Both of these flows are likely to occur in the inlet of a high-speed air-breathing propulsion system. The present research effort is to develop the numerical capability to better predict the performance of an inlet and to better understand fluid dynamical mechanisms associated with the unsteadiness of the flow, which is closely related to the unstart of an inlet. In the present study a supersonic incoming uniform stream ( $M_\infty = 1.32$ ) to a monotonically divergent inlet section is investigated. Depending upon the back pressure prescribed, several different flow features may occur. MacCormack's hybrid method is applied to solve the thin-layer form of the Navier-Stokes equations written in generalized curvilinear coordinates. The Wilcox-Rubesin two-equation model is employed to describe the turbulent eddy viscosity. Suitable characteristic boundary conditions are formulated. Unsteadiness of the flow is demonstrated in the computed results, and surface pressure distributions at various time instants are compared with the time-mean experimental data to illustrate the time-varying nature of the flow. The capability of the present computer code to capture accurately the important physics is demonstrated. Suggestions for further research are offered.

### ACKNOWLEDGMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and fruitful summer at the Flight Dynamics Laboratory, Wright-Patterson AFB, Dayton, OH. He would like to acknowledge the laboratory, in particular the Computational Aerodynamics Group, for its hospitality and excellent working environments.

Finally, he would like to thank Dr. Wilbur L. Hankey for suggesting this area of research and for his collaboration and guidance. He would like to acknowledge not only many helpful discussions with Drs. Joseph Shang and Jim Mace, but also much personal assistance from them to make living in Dayton much easier and to feel at home. Much assistance from Mr. Jeff Graham, Steve Scherr, and Jerry Trummer indeed helped the author to sail along smoothly.



## I. INTRODUCTION

Supersonic flight vehicles require propulsion systems that operate efficiently and reliably in a wide range of altitudes as well as speeds. The inlet must be able to match the airflow requirements of the engine in cruising and maneuvering conditions. The goal of optimizing an inlet for a given mission requires an extensive investigation of inlet performance at all conditions. With the rapid advances in developing efficient numerical algorithms and faster computers, computational fluid dynamics offers a reliable tool to help achieve this goal.

Upon the basis of tremendous success and wide acceptance of the earlier purely explicit method,<sup>1</sup> MacCormack developed a more efficient method<sup>2</sup> in which he applied the idea of characteristics in the region where the integration time step is severely reduced. Furthermore, the split parabolic equation is solved implicitly. Consequently, a much greater time step ( $0(10^2)$  x explicit time step) can be taken and greater efficiency is achieved. While this newer method possesses some advantages, it has not been tested on as many flow problems as the previous method. It has, however, shown success in computing steady and unsteady transonic diffuser flows.<sup>3</sup> This method is thus chosen for numerically calculating supersonic inlet flows investigated by Sajben, et al.<sup>4</sup> The thin-layer form of the Navier-Stokes equations is employed to describe the flow since the nature of the interaction between inviscid and viscous effects and the interaction between shock and traveling waves in the bulk of inlet is so complex that all important terms need to be considered. However, the divergence angle of the inlet is mild and the mesh aspect ratio ( $\Delta x/\Delta y$ ) is large so that the thin-layer approximation obtained by neglecting the streamwise derivatives of viscous terms is considered to be appropriate.

A step in obtaining a better design for the propulsion system is to get a better understanding of the fluid dynamics involved. Therefore, a sensible way of synthesizing available knowledge can be formulated. The difficulties involved in the supersonic inlet flowfield generally lie in (1) ramp and lip geometries which determine the initial shock systems and initial flow conditions, (2) the interaction between the boundary layer and the terminal shock, (3) the adverse pressure gradient set up by the divergence of the inlet in the subsonic diffuser behind the terminal shock, and (4) the amplification of a disturbance by the inevitably existing separated region and the sustaining unsteady flow. Each of these itself is a difficult and unique subject and may be investigated separately. They are, however, coupled in a complicated way in the entire inlet system. The numerical study proposed for the present investigation, therefore, attempts to provide some insight of the flows.

## II. OBJECTIVES

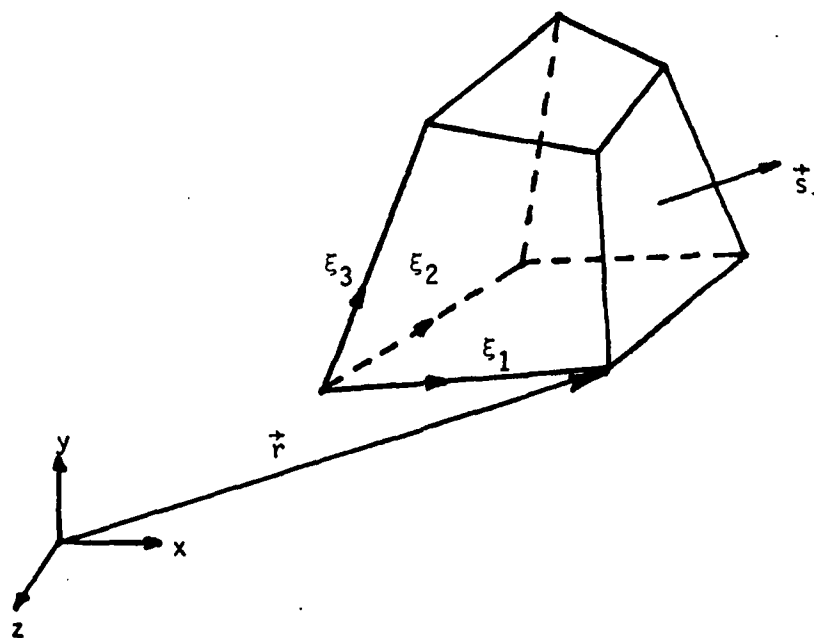
The objective of the research is to develop a technological capability of computing the static and dynamic behavior of the flow in a supersonic inlet that contains all the flow features described above. Work is then directed toward acquiring a sufficient amount of computed data from which the physical mechanism of unsteady flow in an inlet can be extracted. Hopefully this can pave the basis for future research for aiding to identify the unstart margin of an inlet for various geometries and flow conditions.

### III. ANALYSIS

The basic conservation equations and turbulence modeling equations by Wilcox and Rubesin,<sup>5</sup> expressed in vector/dyadic form and in general curvilinear coordinates are:

$$\begin{aligned}
 \partial_t \rho v + \partial_i \rho \vec{u} \cdot \vec{s}_i &= 0 \\
 \partial_t \rho \vec{u} v + \partial_i (\rho \vec{u} \vec{u} + \vec{\sigma}) \cdot \vec{s}_i &= 0 \\
 \partial_t \rho \epsilon v + \partial_i (\rho \epsilon \vec{u} + \vec{u} \cdot \vec{\sigma} + q_e) \cdot \vec{s}_i &= 0 \\
 \partial_t \rho k v + \partial_i (\rho k \vec{u} + \vec{q}_k) \cdot \vec{s}_i &= H_k \\
 \partial_t \rho \omega^2 v + \partial_i (\rho \omega^2 \vec{u} + \vec{q}_\omega) \cdot \vec{s}_i &= H_\omega
 \end{aligned} \tag{1}$$

where  $\partial_t = \partial/\partial t$ ,  $\partial_i = \partial/\partial \xi_i$  are partial derivative operators in time and space, respectively. The  $\xi_i$ ,  $i = 1, 2, 3$ , are curvilinear coordinates related to Cartesian variables through a given transformation. The geometrical variables in Eqs. (1) are  $v$ , the volume element or transformation Jacobian, and  $\vec{s}_i$ ,  $i = 1, 2, 3$ , the surface element vectors, shown in the sketch.



These variables are related to the position vector  $\vec{r}$  through the vector differential formulae:

$$\begin{aligned}\vec{s}_i &= (\partial_j \vec{r}) \times (\partial_k \vec{r}) \\ u &= (\partial_i \vec{r}) \cdot \vec{s}_i\end{aligned}\quad (2)$$

The physical variables appearing in Eqs. (1) are the density  $\rho$ , the total specific energy  $E = e + \vec{u} \cdot \vec{u}/2$ , the specific internal energy  $e$ , the turbulent kinetic energy  $k$ , and a turbulent length scale-defining variable  $\omega^2$ .

The total stress tensor is expressed as

$$\vec{\sigma} = p' \vec{I} - \mu_v u^{-1} [\vec{s}_j (\partial_j \vec{u}) + (\partial_j \vec{u}) \vec{s}_j - \frac{2}{3} \vec{s}_j \cdot \partial_j \vec{u} \vec{I}] \quad (3)$$

where  $\mu_v = \mu + \mu_T$  is a combined viscosity that includes the molecular viscosity  $\mu$  represented by Sutherland's law, and the turbulent eddy viscosity  $\mu_T$ , described by the quantities  $k$  and  $\omega^2$  obtained by solving the differential equations in Eqs. (1). Details of turbulence modeling using the  $k$  and  $\omega^2$  variables can be seen in Ref. 5. The combined pressure  $p'$  is the sum of the thermodynamic static pressure and the dynamic pressure due to turbulence fluctuations  $2/3 \rho k$ .

Analogous to the stress tensor, the thermal and turbulent fluxes are expressed in terms of combined diffusion coefficients

$$\begin{aligned}\vec{q}_e &= -\mu_e u^{-1} \vec{s}_j \partial_j e \\ \vec{q}_k &= -\mu_k u^{-1} \vec{s}_j \partial_j k \\ \vec{q}_\omega &= -\mu_\omega u^{-1} \vec{s}_j \partial_j \omega^2\end{aligned}\quad (4)$$

where  $\mu_e$ ,  $\mu_k$ , and  $\mu_\omega$  are also described by molecular and turbulent eddy viscosity.<sup>5</sup>

The complete system expressed in Eqs. (1) gives the three-dimensional, Reynolds-averaged full Navier-Stokes equations. It is simplified by making the

assumptions: (1) the flow is two-dimensional and (2) the streamwise derivatives in Eqs. (3) and (4) are neglected, the so-called thin-layer approximation.<sup>6</sup> Further, a perfect gas with constant specific heats is considered.

MacCormack's hybrid method is applied to solve the approximate form of Eqs. (1). The numerical procedure discretizes the differential equation using finite-volume concept and time-splits the equations according to the coordinates and types of terms in the equations. That is, the solution vector  $U$ , at a new time level  $t + \Delta t$ , is obtained through a sequence of operators:

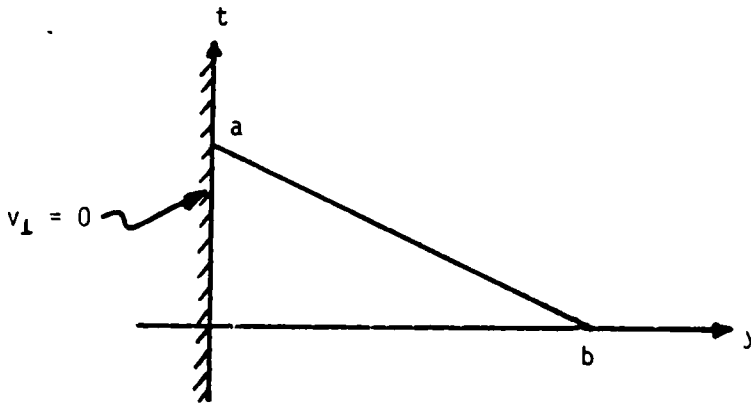
$$U(t+\Delta t) = (L^V(\Delta t) L^{iv}(\Delta t))_{\eta} (L^{iv}(\Delta t))_{\xi} \quad (5)$$

The operators  $L_{\xi}^{iv}$  and  $L_{\eta}^{iv}$  solve the inviscid equations in  $\xi$ - and  $\eta$ -coordinates, and the operator  $L_{\eta}^V$  solves the residual, thin-layer equations in the  $\eta$ -direction.

As the  $L_{\eta}^{iv}$  operator is applied in a fine mesh region close to the solid surface, the space-averaged characteristic relations are used, in place of the earlier explicit operator by MacCormack,<sup>1</sup> to provide intermediate values for the pressure and the velocity component in the  $\eta$ -direction. Subsequent application of MacCormack's explicit operator is then made. Consequently, the maximum stable time step is greatly increased, on the order of 100 times in the case presently considered.

The boundary conditions used in the numerical integration very often play a key role in determining either success or failure. We chose to use the boundary conditions that are based upon the types of equation and flow conditions (physics) involved. The conditions imposed on the viscous operator  $L_{\eta}^V$  are the no-slip, adiabatic, and zero-turbulence conditions at

the wall. The inviscid operators however are imposed by the characteristic equations relating the velocity components and the pressure. Hence, at the solid wall where the normal velocity is zero and the pressure there can be related to quantities at interior points as depicted in the sketch below.



That is,

$$p_a^{n+1} = p_b^n - (\rho c)_{ab}^n v_b^n \quad (6)$$

As for the  $L_{\xi}^{iv}$  operator, the boundary condition becomes more complex and many formulations are proposed, in particular, for subsonic flow boundaries. The inviscid characteristic relations are again adopted. For the supersonic inflow, the domain of dependence of the boundary points lie entirely upstream of it; therefore all variables are prescribed at boundary. If the outflow is subsonic, the  $L_{\xi}^{iv}$  operator possesses two families of characteristic curves (i.e.,  $C^+$  and  $C^0$  in Fig. 1) reaching from the interior of the domain considered and one characteristic curve ( $C^-$ ) from outside. In other words, the domain of dependence associated with the  $C^-$ -characteristic is downstream of the inlet. Hence one variable is allowed to be imposed at the boundary. For a low Mach number boundary, the fixed pressure condition is taken and the other variables are calculated from  $C^0$  and  $C^+$ .

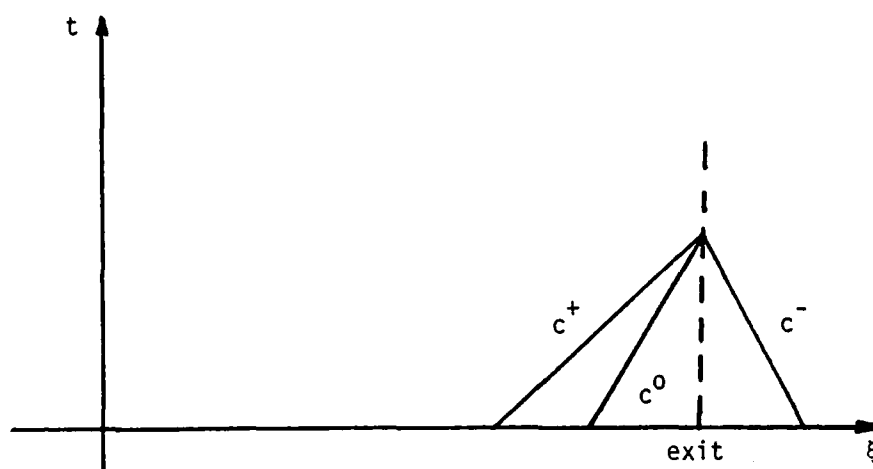


Figure 1. Characteristic curves at the subsonic outflow boundary.

The computational grid systems were chosen so that two grid points were inside the initial sublayer ( $y^+ < 5$ ). The mesh distribution is shown in Fig. 2; eighty cells were used in the streamwise direction and fifty cells were placed across the flow. Uniform, fine meshes were clustered near the shock whose location was indicated by experiment.

#### IV. RESULTS

The problem considered is a two-dimensional inlet with supersonic free stream Mach number of 1.84, shown in Fig. 3. The ramp angle is chosen so that the oblique shock just touches upon the cowl lip ( $x = 3.03$  cm). Since the ramp surface is flat up to  $x = 3.88$  cm, a uniform supersonic flow is generated behind the oblique shock with Mach number 1.32. Therefore the computation domain is chosen to consist of only the flow internal to ramp and lip, indicated by the dotted line in Fig. 3. To determine the validity of the numerical model and the computer code, which is an extended version of the original code developed by Coakley<sup>7</sup> at NASA-Ames Research Center, we chose

to compute the supercritical case<sup>4</sup> with pressure ratio  $PR = 0.724$ , where  $PR$  is defined as the ratio of the static pressure at exit to the stagnation pressure in the plenum. For this case the experiment shows that the terminal shock stays inside the inlet and oscillates, displaying broadband spectral character.

Figure 4a shows the Mach-number contour at 3,300 time steps (approximately equal to the time for a sound wave travelling four (4) inlet lengths) from the initial guess. Two separated flow regions on the top wall and one on the bottom wall are observed. The lambda ( $\lambda$ ) shock pattern is clearly illustrated on both walls; the blown-up picture near the shock is shown in Fig. 4b. A shear layer developing at the foot of the shock is clearly displayed and is convected further downstream. The density contour plot in Fig. 4c shows an expansion and a compression region on the top and bottom walls respectively due to the curvature effect.

Since the computed results indicate that the flow is unsteady in the first two cycles, further Fourier analysis will be of value when more cycles of computed data are made available. Therefore, we show only a sequence of time-evolution of the flow, in comparison with the time-mean data of the experiment by Sajben et al.<sup>4</sup> Figure 5 compares the predicted bottom wall pressure distribution with the experiment. The computed solutions at each instant generally follow the trend of time-mean data. It is worthwhile to note that a compression near the cowl lip is captured and is followed by a rapid expansion till the shock. There also exists an expansion immediately behind the shock/boundary-layer interaction region which is, however, contrary to the phenomena generally observed in the flow over a convex surface, e.g., the upper surface of an airfoil or top wall in our case. We also observe that a large amplitude wave is developing as the shock is moving upstream.



Time variations of the top-wall pressure distribution are shown in Fig. 6; a similar behavior, but different in local features, to the bottom-wall pressure is found.

#### V. RECOMMENDATIONS

At this point, the computed results are by no means complete or conclusive. Yet the following recommendations may be made:

(1) The computer code currently developed has proved to be reliable and appears capable of predicting the unsteady phenomena observed in the experiment. Many fine features of the flow have been correctly predicted. Therefore I propose to continue the research in the following direction.

(2) In order to better understand and quantify the unstart problem so vital to the supersonic propulsion system, there is a need of continuing the present research so that a sufficient amount of data can be made available for providing information of significance scientifically.

(3) Based upon the research carried out in (2), then the unstart problem may be attacked with an objective to determine the margin of inlet instability. With the guidance derived from the close collaboration with analytical work by Adamson, et al at The University of Michigan, we will be able to bracket the amplitude and frequency required for a disturbance to unstart a given inlet.

## REFERENCES

1. MacCormack, R.W., "The Effect of Viscosity on Hypervelocity Impact Cratering," AIAA paper No. 69-0354, 1969.
2. MacCormack, R.W., "An Efficient Numerical Method for Solving the Time-Dependent Compressible Navier-Stokes Equations at High Reynolds Number," Computing in Applied Mechanics, ASME, AMD Vol. 18, 1976.
3. Liou, M-S. and Coakley, T.J., "Numerical Simulations of Unsteady Transonic Flow in Diffusers," AIAA Journal, Vol. 22, pp. 1139-1145, 1984.
4. Sajben, M., Bogar, T.J., Kroutil, J.C., "Experimental Study of Flows in a Two-Dimensional Inlet Model," AIAA paper No. 83-0176, 1983; to be published in AIAA Journal.
5. Wilcox, D.C. and Rubesin, M.W., "Progress in Turbulence Modeling for Complex Flow Fields Including Effects of Compressibility," NASA TP 1517, 1980.
6. Baldwin, G.S. and Lomax, H., "Thin-Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA paper No. 78-257, 1978.
7. Coakley, T.J. and Bergman, M.Y., "Effects of Turbulent Model Selection on the Prediction of Complex Aerodynamic Flows," AIAA paper No. 79-0070, 1979.

# MESH

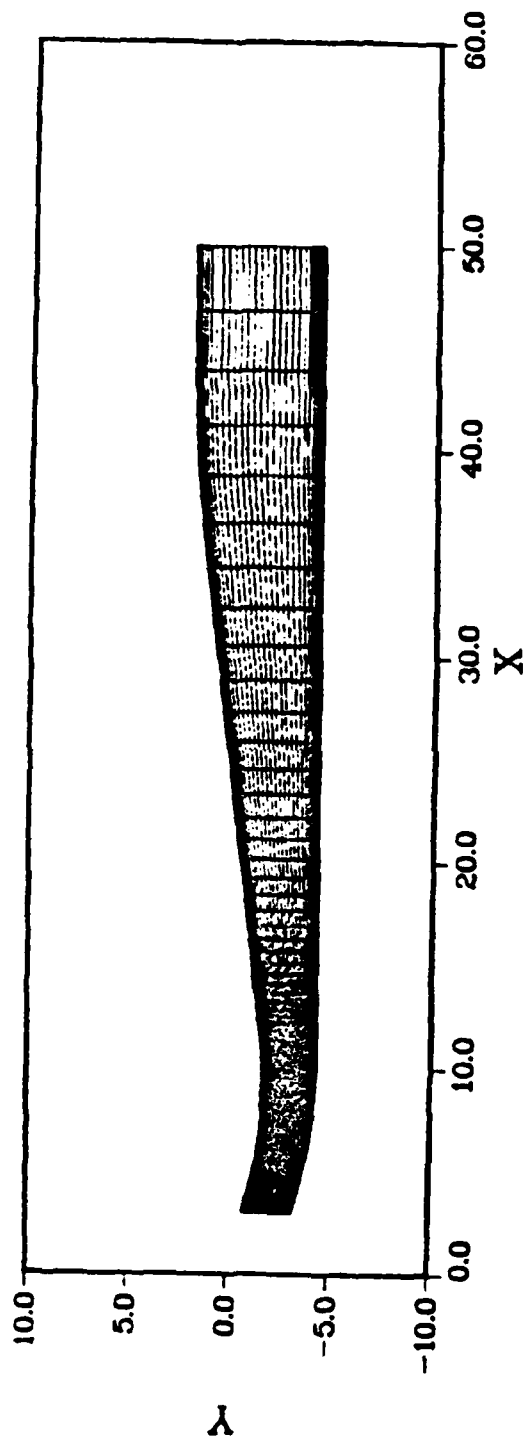


Figure 2. Mesh distribution.

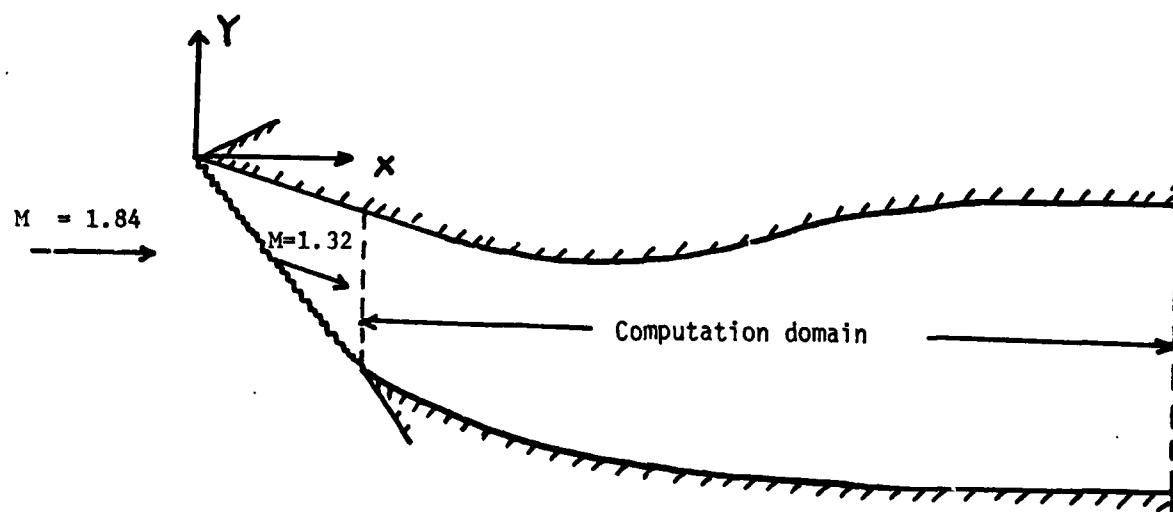
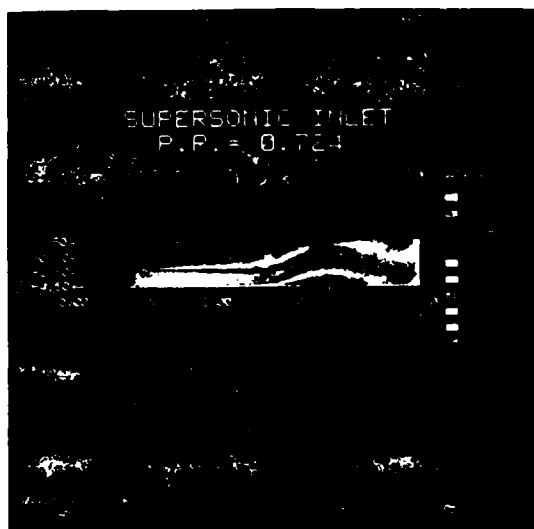
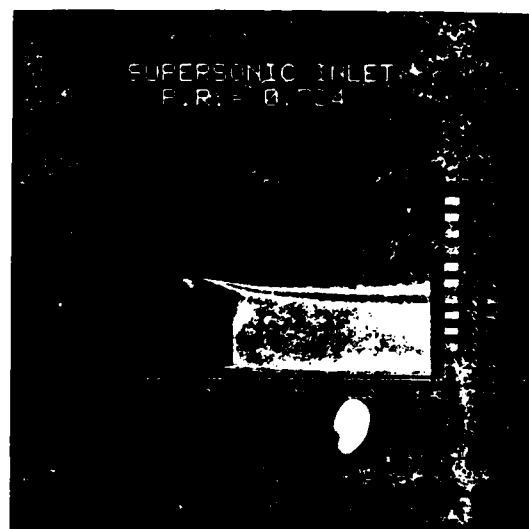


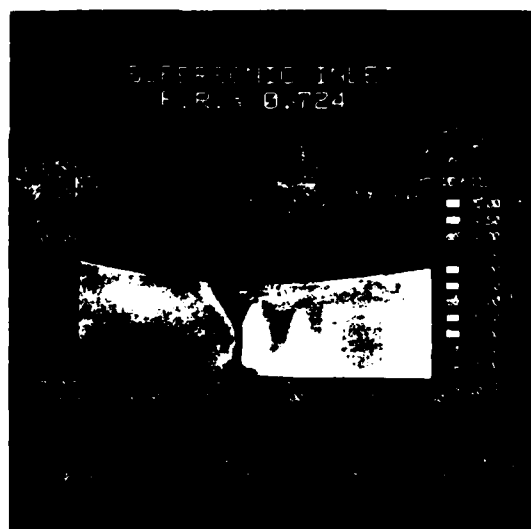
Figure 3. Inlet model .



(a)



(b)



(c)

Figure 4. (a) Mach number contour  
(b) Mach-number contour in blown-up scale near the shock  
(c) Density contour

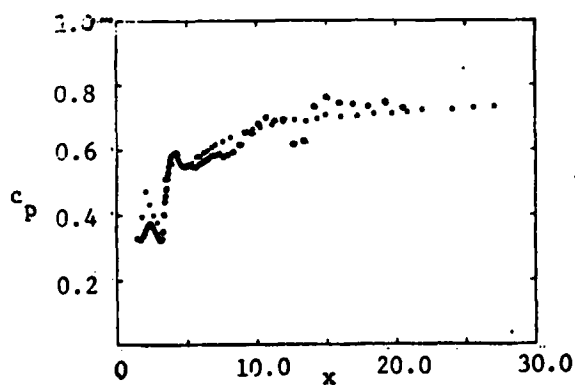
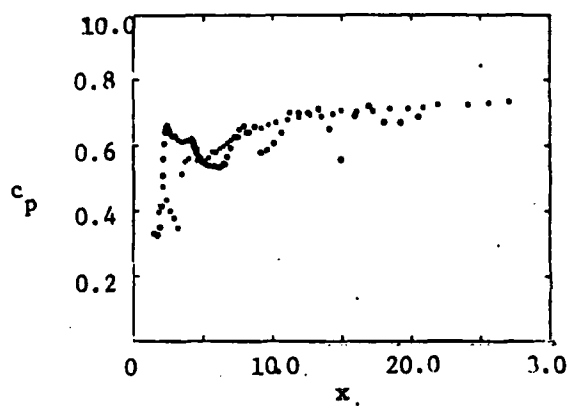
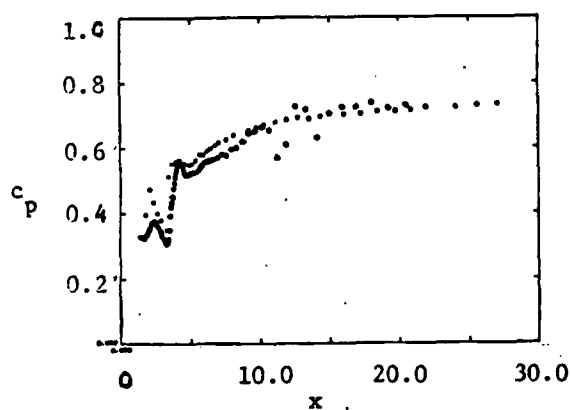
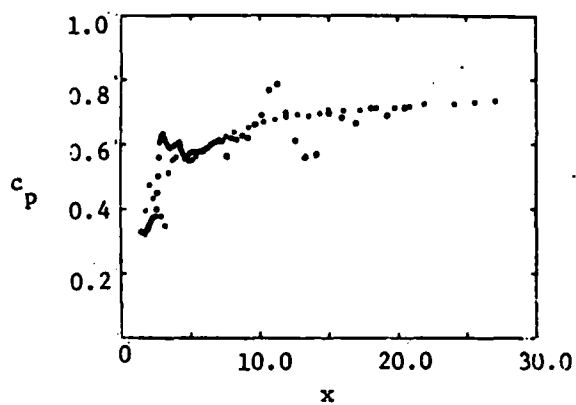
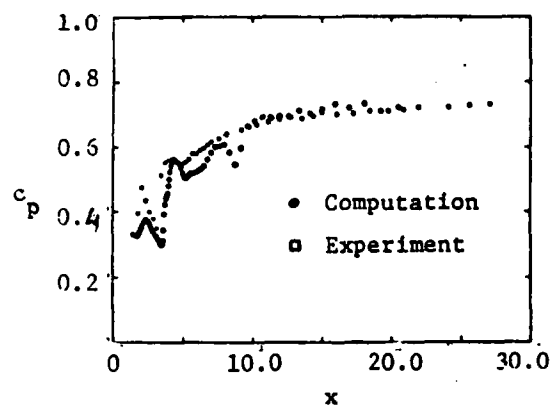


Figure 5. Time-variation of  
bottom-wall pressure distribution  
(every  $1.08 \times 10^{-3}$  sec. interval)

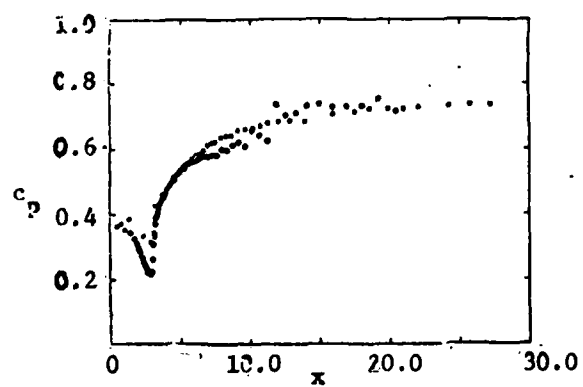
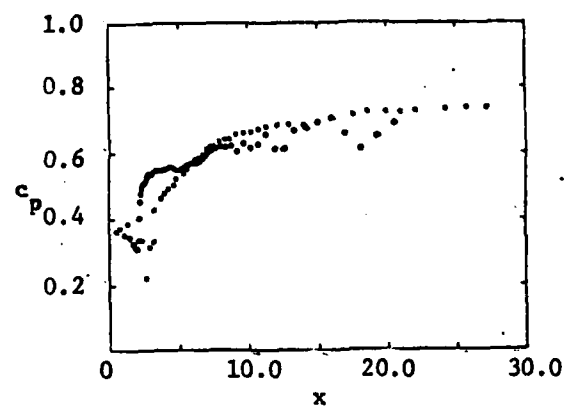
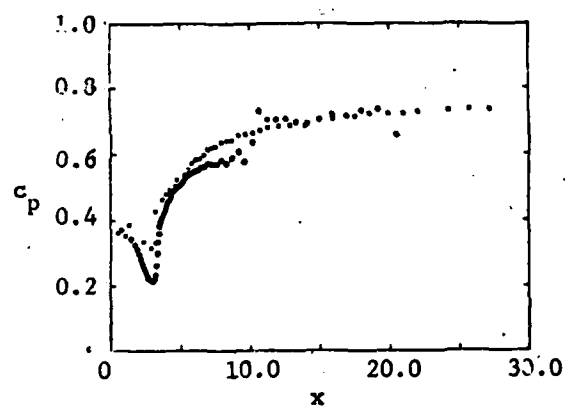
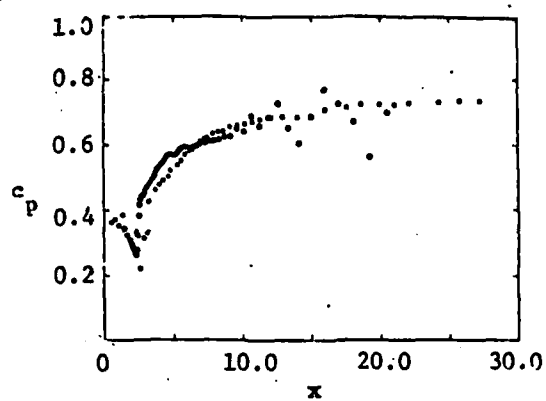
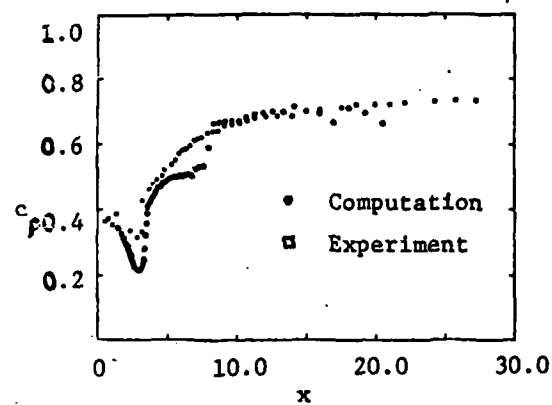


Figure 6. Time-variation of top-wall pressure distribution (every  $1.08 \times 10^{-3}$  sec interval ).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

INFORMATION PROCESSING ANALYSES OF SPATIAL SYNTHESIS AND  
OF THE RELATIONSHIP BETWEEN LEARNING AND INTELLIGENCE

Prepared by:	Dr. David F. Lohman
Academic Rank:	Assistant Professor
Department and University:	College of Education The University of Iowa
Research Location:	Air Force Human Resources Laboratory  Manpower and Personnel Division Test and Training Research Branch
USAF Research:	Dr. Patrick Kyllonen
Date:	September 27, 1984
Contract No.:	F49620-82-C-0035



INFORMATION PROCESSING ANALYSES OF SPATIAL SYNTHESIS AND  
OF THE RELATIONSHIP BETWEEN LEARNING AND INTELLIGENCE

by

David F. Lohman

ABSTRACT

This report summarizes three experiments on the process of mentally synthesizing geometric shapes and two experiments on the relationship between learning and intelligence. The background, method, and preliminary results of each experiment are presented. Experiments on spatial synthesis showed substantial effects for mode of presentation (simultaneous or successive), feedback, and practice. Experiments on learning showed substantial effects in subjects' ability to recognize changes in context and changes in meaning. Implications for information processing theories of spatial ability and intelligence are discussed.

### Acknowledgement

I thank the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the opportunity to collect more data in one summer than I was able to collect in the preceding seven years. I am particularly grateful to Drs. Ray Christal, Pat Kyllonen, and Bill Tirre for their willingness to share the resources of a truly remarkable lab with me, and for their continuing help and support in what has turned out to be a difficult, but enormously worthwhile adventure. I thank Paul Nichols for his help, especially during the first half of our visit when he learned PASCAL and spent many days and nights preparing the stimuli. Thanks also to Tirre's Troups, especially Rich, Ernie, Janet, and Janis. Finally, I thank Mrs. Pamela Wright for the most competent secretarial assistance I have encountered anywhere.

## I. INTRODUCTION

Spatial and mechanical abilities are the best predictors of success in many Air Force training programs, such as those for pilots, air traffic controllers, and aircraft and vehicle mechanics (Anderson, Fruchter, Manuel, and Worchel, 1954). Yet existing paper and pencil tests are often poor measures of spatial ability (Lohman & Kyllonen, 1982). One purpose of this research was to understand better the nature of spatial ability, particularly the process of image construction. The second major purpose of the research was to explore some hypotheses about relationships between measures of general scholastic aptitude and learning. In particular, we sought to test the hypothesis that subjects differ in the amount and type of elaboration they routinely perform when learning and that these differences are systematically related to scholastic aptitude or "intelligence."

## II. OBJECTIVES

Our primary objective was to use to the fullest the data gathering capabilities of the LAMP testing facility at Lackland AFB. Thus, we designed seven different experiments that would be conducted simultaneously. Most of the experiments included multiple conditions with the expectation that one or more conditions might not work as expected. This was done because it was impossible to set up a pilot study, analyze the data, and still have time to conduct the full experiment in the time available. Subjects and testing stations were more plentiful than time for data collection. Further, we realized that data analysis would take months, perhaps years to complete, but that these analyses could be conducted after the fellowship period.

Of the seven experiments that we conducted, five are described in this report and two are described in a companion report by Mr. Paul Nichols. In all cases, we report preliminary results from the first 30 subjects to participate in the experiments. Analyses of the final data sets are presently being conducted both at HRL and at the University of Iowa, and will continue for many months.

### III. EXPERIMENTS ON THE SYNTHESIS OF GEOMETRIC SHAPES

We conducted three experiments on the process of synthesizing geometric shapes. Stimuli used in a previous experiment (Lohman, 1984) were modified for use in these experiments. I first describe the background and research hypotheses for each experiment and then briefly describe the materials and procedures for all three experiments.

Experiment 1: Effects of simultaneous presentation and foil type on mental synthesis.

The first experimental task was adapted from a similar task used by Lohman (1984). The subjects were shown two stimuli (called "A" and "B"), which they were to combine mentally. A test figure (called "C") was then presented. Subjects were to indicate whether C was the correct synthesis of A and B. In contrast with the procedure of Lohman (1979; 1984), both the A and B stimuli were visible at the same time. In previous experiments, the A and B stimuli were presented successively. Simultaneous presentation eliminates the need to memorize the A stimulus, and thereby substantially reduces the difficulty of the task.

#### Effects of Foils

The difficulty of the task is also a function of the difficulty of the foils. In Lohman (1979), foils were created by reflecting the correct C figure about the vertical axis or by arbitrarily deleting some part of the correct C figure. Foils in Lohman (1984) were created by merely altering the proportions of the correct C figure by 15, 25, or 35%. These foils were generally quite similar to the correct C figure, and thus substantially increased the difficulty of the task.

Type of foil was systematically manipulated in the present experiment for three reasons. First, we sought to reduce errors by increasing the discrepancy between foils and correct answers. Second, we hoped to discourage subjects from using strategies that might allow them to respond correctly without having actually synthesized the separate parts. For example, if all foils were constructed by merely reflecting the correct answer about the vertical axis, subjects could respond correctly by remembering only the A

figure and then examining the left side of the test figure. Third, we wanted to use information on the type of errors subjects made to test hypotheses about the nature of spatial ability.

The first hypothesis comes from Smith (1964, p. 96) who argues that the essence of spatial ability is the ability to generate and maintain an image that preserves the correct proportions among the elements of the figure. According to this hypothesis, then, errors on items constructed by altering the proportions of the correct AB figure should show the steepest regression on spatial ability.

Kyllonen (1983) argues that the majority of the many different types of errors he observed on a paper-folding task could be accounted for by a rather simple forgetting function in which information is randomly deleted from a visual short-term memory. According to this second hypothesis, then, false alarms to foils constructed by randomly deleting parts of figures should show the steepest regression on spatial ability, or at least as steep a slope as either of the other two types of errors.

Third, a number of theorists (e.g., Reed, 1974; Palmer, 1977) have hypothesized that figural information is represented in hierarchically organized memory structures called structural units (Palmer, 1975) or structural descriptions (Reed, 1974). Attributes (such as size, orientation, color, etc.) that describe the overall configuration (here, the AB figure) are stored at the highest level in the hierarchy. Attributes of component parts (here, attributes of the separate A and B figures) are stored at a lower level, and attributes of the elements (here, the lines and angles that make up the separate figures) are stored at the lowest level in the hierarchy. Structural descriptions of the correct AB figure and its reflection differ only in the value of a single orientation attribute attached to the highest level unit in the structural description. Thus, subjects who have actually synthesized the separate A and B stimuli into a new C stimulus that has structural properties which cannot be predicted from the separate A and B figures should be more likely to false alarm to foils that are reflections of the correct probe

than to foils in which the structure has been altered.

#### Effects of task difficulty on solution strategy

Previous studies of mental synthesis have shown that, although synthesis tasks show high correlations with other measures of spatial ability, subjects attempt to solve synthesis items using a variety of different strategies (Kyllonen, Lohman, & Woltz, in press; Lohman, 1984). There is some evidence that subjects attempt to solve other complex spatial and figural reasoning tasks by different methods (Bethell-Fox, Lohman & Snow, in press; Lohman & Kyllonen, 1982), but fewer alternative solution strategies have been claimed for these tasks than for the figural synthesis task studied by Kyllonen et al. (in press). Diversity in solution strategy has important implications for the interpretation of test scores, especially for those tests assumed to be unidimensional since variation in solution strategy means that such tests measure different abilities in different subjects, or even different abilities in the same subject on different items (see French, 1965; Lohman & Kyllonen, 1982).

However, previous studies of mental synthesis have used items that vary widely, perhaps too widely, in difficulty. Variations in solution strategy observed by Kyllonen et al. (in press) may have merely reflected the variations in item demands, especially as items became increasingly difficult. Therefore, one purpose of the first experiment was to simplify, or better, to uncomplicate the synthesis tasks used in previous experiments and to determine if subjects still attempted to solve different items using different strategies.

Response latencies are ambiguous for those items which a subject answers incorrectly. Simplifying the task would also simplify the task of modeling response latencies since errors would be reduced. Thus, task simplification would have a collateral benefit of simplifying data analysis.

Thus, the hypotheses tested in the first experiment were:

1. Presenting the A and B figures simultaneously and increasing the disparity between targets and foils will substantially reduce the average difficulty of the items.

2. With easier items, positive correlations between solution latencies and verbal abilities should decrease or disappear altogether while negative correlations between solution latencies and spatial abilities should increase.
3. A decrease in task difficulty should also decrease the number of subjects attempting to solve items by non-spatial strategies. Solution strategy will be estimated by regressing solution latencies and errors on independent variables that represent the amount or duration of a given processing step in an information processing model of task performance (see, e.g., Sternberg, 1977).
4. Correlations between number of errors for each of the three types of foils and reference spatial tests will differ significantly. We have no prediction as to which type of error will show the highest correlation with spatial ability.

#### Experiment 2: Effects of Feedback on Mental Synthesis

Knowledge of results is probably the single most important variable in learning. Learning new responses without feedback of some sort is difficult or impossible (Thorndike, 1932). On the one hand, many experimenters routinely give feedback to subjects, either after each trial or after a block of trials. On the other hand, mental tests are usually administered without feedback. Examiners are admonished to do or say nothing that might reveal to examinees the correctness of their responses (for noteworthy exceptions, see Vygotsky, 1962, and Brown & French, 1979). This is a reasonable practice if one hopes to hold the object of measurement steady while one measures it. Furthermore, it probably enhances the predictive validity of scholastic aptitude tests since most instruction does not provide immediate feedback either (see Snow & Lohman, 1984). Therefore, one purpose of the present experiment was to compare the experimental task procedure of providing feedback with the mental test procedure of withholding feedback on a test-like task.

Second, feedback has been used in a effort to train spatial abilities. Levine et al. (1979) gave subjects feedback on correctness while they practiced spatial test items like those found on standard spatial tests. Levine et al. (1979) found that although subjects generally improved with practice, the effects did not transfer to another spatial test. However, feedback consisted merely of response correctness. We planned to allow subjects the opportunity to determine how they missed an item by showing the full item (stem, foil, and correct answer). Simple feedback on response correctness should suffice only for those who are already aware of their mistakes or whose mistakes are due to haste or carelessness.

Kyllonen, Lohman, and Snow (1984) studied the effects of a brief training on subjects' ability to solve paper-folding items. They exposed subjects to one of three treatments: visual modeling, verbal labeling, or practice with feedback. The simple practice with feedback was found to be the most effective treatment for subjects high in spatial ability, but not for other subjects. However, treatments were quite short, and there was no way to assess whether all subjects studied the feedback. We planned to provide explicit feedback and to record how long subjects studied it. Thus, the present study extends the work of Levine et al. (1979) by providing feedback that consists of more than response correctness and to extend the work of Kyllonen et al. (1984) by providing an estimate of how carefully the subjects attend to the feedback that is provided.

Specifically, our predictions were: (1) Subjects receiving feedback will show more improvement with practice than subjects not receiving feedback. (2) In contrast to the results of Kyllonen, Lohman and Snow (1984), the benefits of feedback should be greatest for the least able subjects and least for the most able subjects. (3) Effects of feedback will be moderated by the amount of time subjects spend studying the feedback provided. Those who study the feedback more will show the greatest benefit. (4) Feedback should accelerate the effects of practice. Thus, the regression of solution latency on spatial ability should be steeper in the feedback condition than in the no-feedback condition.



Experiment 3: Effects of extended practice without feedback on the relationships between synthesis task latencies, errors, and reference spatial and verbal abilities.

The purpose of the third experiment was to examine the changes in the correlations between latencies and errors on the experimental synthesis task and reference abilities as subjects were given extended practice. In brief, our hypothesis was that latencies (and slopes derived from them) would initially show low negative correlations with reference spatial tests and might even correlate positively with verbal reference tests. With practice, however, we expected that the negative correlations between spatial tests and latencies would increase. Errors were expected to show the opposite pattern: highest negative correlations with spatial ability should occur early and then decrease with practice.

#### Method

##### Subjects and Reference Tests

Subjects were approximately 1000 Air Force enlistees selected randomly from all Air Force enlistees on their sixth day of basic training. From two to four months prior to the experimental testing session, subjects had taken Version 9 or 10 of the Armed Services Vocational Aptitude Battery (ASVAB). In addition to the ASVAB, four paper-and-pencil tests were administered immediately prior to the experimental session. The tests were: Identical Pictures (French et al., 1963), Part 2 of Form Board (French et al., 1963), Block Synthesis (HRL Test files), and Gestalt Completion (French et al., 1963 and HRL Test files). On all tests, subjects marked directly on the tests booklet since we did not wish to confound time to answer items with time to find the answer box on a machine scorable answer sheet.

##### Experimental Task

The experimental task consisted of two blocks of 72 synthesis items. The difficulty of combining forms is primarily a function of the complexity of the separate pieces and the complexity of the product image (Kyllonen, Lohman, & Woltz, in press; Lohman, 1979). Therefore, items varied orthogonally on the average complexity of the A and B figures (4-5, 6-7, or 8-9 points), and the correctness of the test probe C.

Six replicates of each item type were constructed yielding 3x2x2x6 or 72 items in all. Foils were then constructed for the three negative trial replicates either by reflecting the correct AB figure about the vertical axis, by deleting part of the correct AB figure, or by altering the proportions of the correct AB figure by 35%. Two new blocks of items were created temporarily by reflecting the items in Blocks 1 and 2 about the horizontal axis. The final two blocks of items were then formed by assigning the upright or reflected version of each stimulus either to Block 1 or Block 2. Thus, each of the final two blocks of items contained an equivalent number of upright and reflected stimuli. Items in Block 1 were physically identical but inverted versions of items in Block 2. Items were then ordered so that the main effects for each design facet were balanced within groups of six items. These 18, six-item blocks were presented in the same order to all subjects. However, the six items within each block were randomly ordered for each subject.

#### Procedure

Subjects were randomly assigned to one of three conditions. Subjects in Group 1 (No Feedback) received no feedback after each trial. Subjects in Group 2 (Feedback) received a message indicating the correctness of the response and the time taken to accept or reject the C figure after each trial. For incorrect responses, subjects in the Feedback condition were shown the full item again (A, B, and C figures). If the subject had erroneously accepted an incorrect C figure, then the correct AB figure was shown below the incorrect C figure on the screen. If the subject had erroneously rejected a correct C figure then the full item was shown again along with a message indicating that the probe was in fact correct. These figures remained on the screen until subjects pressed a key to continue.

Subjects in Group 3 (Repetition) were administered the task three times without feedback. They were given a short break between successive presentations and were told that some items would be repeated the second and third times they attempted the task.

#### Results

Preliminary results reported here are based on a sample of 13

subjects who participated in the experiment on the first day. Subjects who received feedback made fewer errors than subjects not receiving feedback. Further, differences between the two groups increased with practice. Subjects in the feedback condition answered 73% of the items in Block 1 correctly and 80% of the items in Block 2 correctly. Corresponding means for subjects not receiving feedback were 67% and 71%. Thus, although all subjects improved with practice, those receiving feedback improved most.

Feedback had much smaller effects on latencies. The only noticeable difference was that subjects not receiving feedback spent less time attempting to synthesize the A and B figures. However, subjects in both groups accepted or rejected the test probe C with equal speed. Also, subjects in the feedback condition did appear to be studying the feedback stimuli, and not only in the first block. Mean time on feedback frames was 3.97 and 3.50 seconds for Blocks 1 and 2, respectively.

The only negative result in these data was that subjects in the repetition group appeared not to be improving with practice. Mean correctness dropped from 70% on the second presentation of the items to 65% on the third attempt. However, subjects were not told to expect repetitions of items and some later reported that they thought that something was wrong with their terminal. We modified the instructions and so results may differ for the full sample.

#### IV. EXPERIMENTS ON THE RELATIONSHIP BETWEEN LEARNING AND INTELLIGENCE

##### Experiment 4: The Role of Semantic Elaboration and Contextual Cues on the Relationships Between Learning and Intelligence

Although intelligence and learning are two of the most important constructs in education, we understand little about the relationships between them. Early theorists presumed that intelligence and the ability to learn were synonymous (see, e.g., McGeoch, 1952) in part because of the consistently high correlations observed between measures of intelligence and achievement (see Jensen, 1980 for a review). But ability-learning correlations, although significant, have generally been low, and have not supported a general learning factor (see Cronbach and Snow, 1977; Gagne, 1967; and Snow, Kyllonen, &

Marshall, 1984). Thus, the paradox is that intelligence is a good predictor of scholastic achievement, but appears to be a poor predictor of learning.

A recent study reported by Sorensen (1983) suggests that the relationship between ability and learning is moderated by the amount of inferencing or elaboration performed by the learner. The Sorensen (1983) study suggested that more able students elaborate more relationships (and perhaps more automatically) than less able students. It is quite possible, however, that less able students do not elaborate less, just differently. In particular, general ability may indicate a tendency to decontextualize knowledge. High G subjects may habitually elaborate semantically whereas low G subjects elaborate contextually. The tendency to store information by elaborating to contextual cues is disfunctional in school for several reasons. First, after a while, one school day is much like the next, and so it becomes increasingly difficult to retrieve information using contextual cues. Second, cues given in tests are almost invariably semantic, and thus favor those persons who have stored (or, better, elaborated) information semantically. Episodic (i.e., context-bound) memories tend to be personal and idiosyncratic and are not as easily activated by a common cue. The fourth experiment was designed to test this hypothesis.

#### Subjects

Subjects were approximately 1000 Air Force enlistees selected randomly from all Air Force enlistees on their sixth day of basic training.

#### Materials

Twenty-four sentences containing a homograph were constructed. Homographs were taken from the list compiled by Nelson, McEvoy, Walling, and Wheeler (1980). Two matching foil sentences were then constructed for each target sentence. The first type of foil contained virtually the same words as the original sentence but capitalized on the alternate meaning of the homograph. For example, the target sentence "Jim struck Bob with a club" became "Jim struck Bob at the club." We

expected that subjects who remembered the meaning of the target sentence rather than the exact wording of the sentence would be most likely to false alarm to these sentences. We refer to these three types of sentences as the target sentences, the homograph foils, and the synonym foils, respectively. Lastly, 24 nonsense sentences were created.

The incidental learning pretest was then created by randomly combining the 24 target sentences and the 24 nonsense sentences in a single 48-sentence list under the constraint that no more than three nonsense sentences or three target sentences occur successively. Half of the subjects saw these 48 sentences with contextual cues whereas half of the subjects saw all sentences in normal type in the center of the CRT. Twenty-four different combinations of contextual cues were created by presenting a sentence in one of six different locations on the CRT (top left, middle left, bottom left, top right, middle right, or bottom right), in one of two type faces (standard or script), and in one of two modes (regular or inverse video). One target sentence and one nonsense sentence were assigned to each of the 24 different contexts.

Three 32-item recognition tests were then created. All contained the same set of eight randomly selected nonsense sentences. The remaining 24 items were composed of eight target sentences, eight homograph foils, and eight synonym foils. Thus for each of the 24 target sentences, the target version of the sentence appeared in one form of test, the homograph foil version of the sentence appeared in the second form of test, and the synonym foil version of the test appeared in the third form of the test.

Finally, for subjects in the context condition, half of the sentences on the posttest were presented in the same context in which they had appeared on the pretest. The other half of the sentences appeared in a different context than they had appeared on the pretest.

#### Procedure

Subjects were randomly assigned to one of the six conditions defined by the crossing of context (yes, no) with form of posttest (A, B, or C). They were instructed to read each of the 48 sentences

carefully and then decide whether the sentence was sensible.

On the recognition posttest, subjects were told that they were to decide if each sentence was the same as one of the sentences they saw earlier. Subjects in the context condition were also told that they should press "same" if the sentence was the same as one they had seen earlier even if the sentence appeared in a different context.

### Results

Preliminary results reported here are based on 30 subjects who participated in the experiment on the first day. However, several effects were quite strong, even in such a small sample. First, context made a difference, both on the pretest and on the posttest. Those who saw sentences in context on the pretest had difficulty distinguishing sensible sentences from nonsensical sentences (mean correct 72% and 82% for context and no context groups, respectively). More importantly, context had significant effects on the posttest. Sentences that appeared with the same contextual cues on both pretest and posttest were recognized much better ( $\bar{x}=76.9$ ) than sentences that appeared in a different context on pretest and posttest ( $\bar{x}=71.9$ ).

The second important result was that subjects rejected the homophone foils more often than the synonym foils. This supports many other experiments which show that people are generally more sensitive to meaning changes than to changes in wording or syntax. The important question of whether sensitivity to change in meaning or changes in context is related to ability cannot be answered by these data due to sample size. Nevertheless, these preliminary results are encouraging in that they suggest that the experimental manipulations worked as expected.

Experiment 7: The relationship between ability, study time, and perceived relatedness of noun pairs in an incidental learning task.

Able learners create meaning, even where none was intended. The purpose of the last study was to test the hypothesis that the meaning-seeking habits of high G subjects help them to learn even when they are not trying to learn. In particular, the hypotheses were:

(1) High ability subjects will "see" more relationships between randomly paired nouns than will low ability subjects. (2) The relationship between general ability and the number relationships reported between noun pairs will be stronger for low imagery noun pairs than for high imagery noun pairs. (3) Recognition memory for noun pairs will be significantly related to both (a) the perceived relatedness of nouns and (b) the time spent examining the noun pairs.

#### Design and Materials

This experiment had three phases: pretest, retention, and posttest. In the pretest, 48 noun pairs were presented one at a time. Subjects were to decide if the two words were related in some way. During the retention phase they solved spatial problems for approximately one hour and took a 24-item recognition memory test for an incidental sentence learning task. In the posttest phase, 96 noun pairs were presented one at a time. Subjects were asked to recognize the 48 noun pairs presented in the pretest.

The 48-item pretest contained three types of items. One-third of the pairs (16 items) contained nouns selected randomly from Palermo and Jenkins (1964) to be weakly associated (10 or fewer occurrences in the high school sample). One-third of the pairs (16 items) were constructed by randomly pairing high frequency (A or AA in Thorndike & Lorge, 1944), high imagery nouns (rated imaginability  $\geq 6.0$  in Paivio, Yuelle, & Madigan, 1968). The final one-third of the pairs (16 items) were constructed by randomly pairing high frequency (A or AA in Thorndike & Lorge, 1944), low imagery nouns (rated imaginability  $\leq 3.9$  in Paivio et al., 1968). We refer to these as the weak-associate pairs, the random-high-imagery pairs, and the random-low-imagery pairs. Items were then randomized under the constraint that each quarter of the pretest contain four items from each of the three categories. All subjects judged items in the same random order.

The 96-item posttest contained the 48 pretest word pairs plus 48 foils. Three types of foils were constructed. One-third (16 foils) contained noun pairs that were strong associates. These pairs were constructed by randomly selecting noun pairs from Palermo and Jenkins (1964) under the constraint that both nouns be high in rated imagery

and high frequency. Subjects who judged the familiarity of noun pairs on the strength of the association between terms should be misled by these items. The next 16 foils were long associates (20-101 occurrences in the Palermo & Jenkins high school sample) of a representative sample of noun pairs presented in the pretest. We expected that subjects who noticed relationships between pairs on the pretest would false alarm to these pairs. The final one-third (16 foils) were constructed by randomly pairing low frequency nouns from the Paivio et al. (1968) list. These nouns should not have been familiar in any way to subjects. False alarms here suggest carelessness, random responding, or substantial difficulties in recognition memory. All 96 items were then randomly assigned to four 24-item blocks under the constraint that each block contain an equal number of each of the three types of target items and of the three types of foils. All subjects saw items in the same random order.

#### Procedure

Subjects were told that they should decide whether each pair of words that appeared on the CRT was related in some way. They were told to press the "l" key if the two words were related in some way and to press the "d" key if the two words were not related. Intertrial interval was 750 msec.

During the 45 minute to 1 hour retention interval, subjects solved spatial problems as part of another experiment. They also took a 24-item recognition memory test for sentences they had judged earlier as part of another experiment.

On the posttest, subjects were told to press "l" if the word pair was one presented earlier and to press "d" for new word pairs. They were instructed to read the words carefully but to decide as quickly as possible.

#### Results

Preliminary results reported here are based on the responses of the 30 subjects who participated in the experiment on the first day. Perhaps the most surprising result was that subjects reported seeing a relationship between the random low-imagery word pairs more frequently than for either of the other two types of word pairs. Further, they spent less time reading the random low-imagery pairs than other word



pairs, suggesting that the greater perceived relatedness was not due to longer processing. However, as with the other experiments, the major questions for this experiment cannot be answered until the within-subject analyses are completed for the full sample of recruits.

#### V. RECOMMENDATIONS

Strong recommendations for future research or for practice are not possible until data analyses for each experiment are completed. We do believe, however, that these preliminary results together with the discussions of these sorts of problems with Project LAMP personnel allow the following suggestions. (1) Self-paced experimental tasks produce latencies that are deficient in several respects. Problems of persistence and speed-accuracy trade-off are the most troublesome. Attempts to equate subjects in speed-accuracy trade-off through instructions have not been particularly successful. Nor have signal detection analyses proved very useful, especially since most experimental tasks are not unidimensional and such analyses require enormous quantities of data in most cases. Adaptive pacing of items appears to offer a more reasonable alternative. Items (or parts of items) would be presented for predetermined duration. The dependent variable then becomes probability correct for different exposure durations for items of a given level of complexity. Such schemes for estimating response latency deserve more careful study. (2) Existing measures of mental ability provide only the roughest of guides for the development of new measures of learning ability. To avoid reinventing existing measures, new or improved measures of learning ability must be validated against better measures of learning or performance, not simply correlated with existing mental measures. (3) Faceted tests provide a powerful paradigm for investigating hypotheses about the sources of individual differences in complex tasks. However, when this sort of faceted design is combined with a between-subjects treatment variable and reference ability tests are also administered, then the machinery of ATI analyses can also be brought to bear. If comprehensible, of this marriage of analytic techniques will be truly awesome.

### References

1. G.V. Anderson, B. Fruchter, I.M. Herschel, & P. Worchel, "Survey of Research on Spatial Factors," Research Bulletin 54-84, Air Force Personnel & Training Research Center, Lackland, AFB, 1954.
2. C.E. Bethell-Fox, D.F. Lohman, & R.E. Snow, "Adaptive Reasoning: Componential and Eye Movement Analysis of Geometric Analogy Performance," Intelligence, in press.
3. A.L. Brown & L.A. French "The Zone of Proximal Development: Implications for Intelligence Testing in the Year 2000," Intelligence, Vol. 3, pp. 253-271, 1979.
4. L.J. Cronbach, & R.E. Snow, Aptitudes and Instructional Methods: A Handbook for Research Interactions, (Irvington, New York, 1977).
5. J.W. French, "The Relationship of Problem-solving Styles to the Factor Composition of Tests," Educational and Psychological Measurement, Vol. 15, pp. 9-28, 1965.
6. J.W. French, R.B. Ekstrom, & L.A. Price "Kit of Reference Tests for Cognitive Factors," Princeton, New Jersey: Educational Testing Service, 1963.
7. R.M. Gagne (Ed.), Learning and Individual Differences, (Merill, Columbus, Ohio, 1967).
8. A.R. Jensen, Bias in Mental Testing, (Free Press, New York, 1980).
9. P.C. Kyllonen, "An Information Processing Model of Spatial Ability," Unpublished doctoral dissertation, Stanford University, Stanford, California, 1983.
10. P.C. Kyllonen, D.F. Lohman, & R.E. Snow, "Effects of Task Facets and Strategy Training on Spatial Task Performance," Journal of Educational Psychology, Vol. 76, pp. 130-145, 1984.
11. P.C. Kyllonen, D.F. Lohman, & D. Woltz, "Componential Modeling of Alternative Strategies for Performing Spatial Tasks," Journal of Educational Psychology, in press.
12. J.M. Levine, D. Schulman, R.E. Brahlek, & E.A. Fleishman, "Training Ability of Abilities: Training and Transfer of Spatial Visualization," Washington, D.C., 1979.
13. D.F. Lohman, "Spatial Ability: Individual Differences in Speed and Level," (Tech. Rep. No. 9), Aptitude Research Project, School of Education, Stanford University, 1979.
14. D.F. Lohman, "The Role of Prior Knowledge and Strategy Shifting in Spatial Ability," Paper presented at AERA Meeting, New Orleans, LA, April, 1984.

15. D.F. Lohman, & P.C. Kyllonen, "Individual Differences in Solution Strategy on Spatial Tasks," In R.F. Dillon & R.R. Schmeck (Eds.), Individual Differences in Cognition, (Academic Press, New York, 1983).
16. J.A. McGeoch, The Psychology of Human Learning, (Longmans, Green & Co., New York, 1942).
17. D.L. Nelson, C.L. McEvoy, J.R. Walling, J.W. Wheeler, "The University of South Florida Homograph Norms," Behavior Research Methods and Instrumentation, Vol. 12, pp. 16-37, 1980.
18. A. Paivio, J.C. Yuille, S.A. Madigan, "Concreteness, Imagery and Meaningfulness Values for 925 Nouns," Journal of Experimental Psychology Supplement, Vol. 76, (1, Pt. 2), 1968.
19. D.S. Palermo, & J.J. Jenkins, Word Association Norms, (University of Minnesota Press, Minneapolis, 1964).
20. D.E. Palmer "Hierarchical Structure in Perceptual Representation," Cognitive Psychology, Vol. 9, pp. 441-474, 1977.
21. S.K. Reed "Structural Descriptions of Visual Images," Memory and Cognition, Vol. 2, pp. 329-336, 1974.
22. I.M. Smith, Spatial Ability, (Knapp, San Diego, CA, 1964).
23. R.E. Snow, P.C. Kyllonen, & B. Marshalek, "The Topography of Ability and Learning Correlations," In R.J. Sternberg (Ed.), Advances in the Psychology of Human Intelligence, Vol. 2, (Erlbaum, Hillsdale, N.J., 1984).
24. R.E. Snow & D.F. Lohman, "Toward a Theory of Cognitive Aptitude for Learning from Instruction," Journal of Educational Psychology, Vol. 76, pp. 347-376, 1984.
25. H.B. Sorensen, "The Relationship Between General Ability and Semantic Integration," Unpublished masters thesis, The University of Iowa, College of Education, 1983.
26. R.J. Sternberg, Intelligence, Information Processing, and Analogical Reasoning: The Componential Analysis of Human Abilities, (Erlbaum, Hillsdale, N.J., 1977).
27. A.L. Thompson & R.L. Klatzky, "Studies of Visual Synthesis: Integration of Fragments into Forms," Journal of Experimental Psychology: Human Perception and Performance, Vol. 4, pp. 244-263, 1978.
28. E.L. Thorndike, The Fundamentals of Learning, (Teachers College, New York, 1938).
29. E.L. Thorndike & I. Lorge, The Teacher's Word Book of 30,000 Words, New York: Teachers College Press, 1944.
30. L.S. Vygotsky, Thought and Language, (Wiley, New York, 1962).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

SUBOPTIMUM EXTRAPOLATION FOR SPECTRAL ESTIMATION

Prepared by:	Lonnie C. Ludeman
Academic Rank:	Professor
Department and University:	Electrical and Computer Engineering Dept. New Mexico State University
Research Location:	Rome Air Development Center, Signal Processing Section
USAF Research	Paul Van Etten
Date:	August 15, 1984
Contract No:	F49620-82-C-0035

SUBOPTIMUM EXTRAPOLATION  
FOR SPECTRAL ESTIMATION

by

Lonnie C. Ludeman

ABSTRACT

Several analog systems for extrapolation of a section of analog data were investigated for possible use in spectral estimation. Systems considered had sectionlized and continuous logarithmic modulated and quadratic phase characteristics. The logarithmic phase system had extrapolation characteristics; however, it was very restricted in terms of bandwidth to effective center frequency ratio. Upon sectionalizing, a wider ratio could be obtained; however, the length of useful extrapolation became quite limited.

#### ACKNOWLEDGEMENT

The author would like to thank Paul Van Etten of the Rome Air Development Center (RADC) at Griffiss AFB N.Y. for formulating an interesting area of pursuit and providing guidance, spirited discussions and help when needed. The Signal Processing Section of RADC provided a very fertile atmosphere for my research effort. The acceptance and help given by Russ Brown and all other members of the Signal Processing Section was greatly appreciated and made my stay enjoyable and productive.

I also thank the Air Force System Command and the Air Force Office of Scientific Research for sponsoring my work on the Summer Faculty Research Program and for the able and considerate administration of the program by the Southeastern Center for Electrical Engineering Education.

## I. INTRODUCTION

In radar applications a very important basic problem is that of detection of targets in noise. In many cases this problem can be reduced to the detection and estimation of a small number of sinusoidal signals in noise. Based on the original statistical assumptions and apriori information on the signal and noise, optimum linear and nonlinear detectors and estimators have been developed. These include maximum a posteriori, maximum likelihood, least squares and Bayes estimators and detectors. Most of the methods explored involve solving sets of linear and non linear algebraic and integral equations and an eventual quantization.

Normally a piece of a continuous time signal is available and the signals we are looking for are contained within a given band of frequencies. By sampling the signal a set of values is obtained from which the frequencies and amplitude of the sinusoids can be estimated. Prony [1] developed a simple elegant method for obtaining the coefficients and frequencies of the sum of exponential signals (includes damped sinusoids) and thus provided a means for extrapolating the signals in clean data. The maximum entropy technique developed by Burg [2] provided a way of extrapolating the signal and realized so called "super resolution" by using an autoregressive model. Papoulis [3] developed an iterative algorithm for the extrapolation of data. Recently Bolling [4] has given examples that show improved spectral estimation can be obtained by extrapolating the data using linear prediction and the Fast Fourier Transforms (FFT) of the extended

data. The above methods are in general nonlinear processing of the data and are data dependent. When the data is noisy, data dependence can be degrading. Without modifications the techniques given above perform poorly in noise resulting in significant biases and variance in the estimated values along with poor relative amplitude discrimination.

If the signal is not sampled, the continuous data can be optimally processed for smoothing and extrapolation by use of causal or noncausal Wiener Filters [5] or a Kalman Filter [6]. These filters are determined from the mean and power spectral densities of the signal and noise and changes in spectral densities produce different optimal filters; however, they are not data dependent.

To react to changing environments requires adaptive filtering to maintain a sense of optimality along with the usual problems of time to adapt and stability. A logical question then becomes is there a reasonable non-adaptive non-time varying method for extrapolating a signal so that reasonable results can be obtained under a wide variety of situations.

## II. OBJECTIVES

The main objective of this paper is to explore various suboptimum analog filters for their ability to extrapolate a section of continuous data. This extrapolated data and the original data could then be used to obtain improved spectral estimates.



### III. BACKGROUND

Let  $x(t)$  and  $y(t)$  represent the input and output to a linear system specified by its transfer function  $H(s)$  as shown in Figure 1.

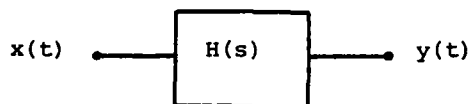


Figure 1. General Linear System

If  $H(s)$  is selected as

$$H(s) = \exp(-\tau_1 s) \quad (1)$$

then the output signal  $y(t)$  is simply a translated version of the input signal, that is

$$y(t) = \mathcal{L}^{-1}[H(s)X(s)] = \mathcal{L}^{-1}[\exp(-\tau_1 s)X(s)] = x(t-\tau_1) \quad (2)$$

Therefore if the input is a stepped sinusoidal signal as shown in Figure 2, then the output will be a delayed stepped sinusoid. The frequency response of such a filter is given by

$$H(j\omega) = \exp(-j\omega\tau_1) \quad (3)$$

and has a corresponding magnitude and phase characteristic  $|H(j\omega)|$  and  $\phi(\omega)$  given by

$$|H(j\omega)| = 1 \quad \phi(\omega) = -\omega\tau_1 \quad (4)$$

If the input to the system is a stepped unit sinusoid of arbitrary radian frequency  $\omega_0$  then the phase difference between the input and output is  $-\omega_0\tau_1$  and the steady state output is given by

$$y(t) = \sin(\omega_0 t - \omega_0\tau_1) = \sin[\omega_0 (t - \tau_1)] \quad (5)$$

For a given  $\omega_0$  it is easy to select a  $\tau_1$  such that the output is out of phase with the input as shown in Figure 2.

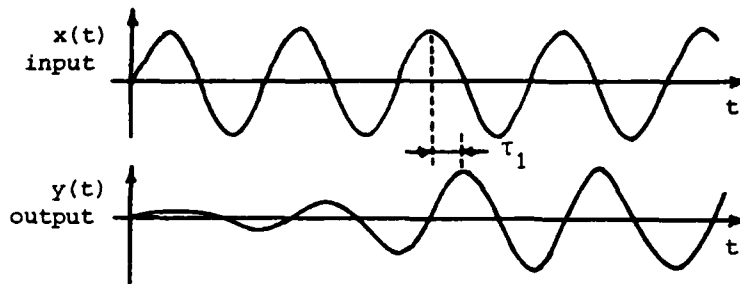


Figure 2 Input and output for linear system with  $H(s) = \exp(-s\tau_1)$ .

For any given  $\tau_1$ , there are many possible  $\omega_0$ 's such that the output is out of phase with the input.

The group delay for an analog filter  $H(s)$  is defined as

$$\tau(\omega) = -d\phi(\omega)/d\omega. \quad (6)$$

It represents the approximate time delay that an analog input signal of radian frequency  $\omega$  is delayed by the filter.

If the phase is linear then  $\tau(\omega)$  will be constant representing a fixed time delay through the system for all frequencies.

For an analog filter to be useful as an extrapolater it must extend the signal without disturbing the frequency content. In this way the phase of each of the sinusoidal components of the extrapolated signal must line up with the corresponding sinusoidal components of the original signal. One way to achieve this is to make the phase zero; however, then the filter has very little extrapolation capability.

#### IV. ANALYSIS

Several systems were investigated for possible use as extrapolators. Among these were the logarithmic modulated phase, piecewise logarithmic phase, and quadratic phase signals. The characteristics of these systems are described in the following sections:

##### A. Logarithmic Modulated Phase Signal

The first system investigated was one that has a different group delay  $\tau_i$  for every input sinusoidal signal of radian frequency  $\omega_i$ . The time delay is chosen to correspond to a fixed multiple  $N$  of periods of a sinusoid input as follows:

$$\tau_i = N \cdot 2\pi / \omega_i = N / f_i \quad (7)$$

In general then the time delay  $\tau_\omega$  (so called group delay) can be written as

$$\tau_\omega = N \cdot 2\pi / \omega \quad (8)$$

and can be approximated by the negative derivative of the phase characteristic of the system. That is

$$-d\phi(\omega)/d\omega = N \cdot 2\pi / \omega \quad (9)$$

Integrating equation (9) the phase  $\phi(\omega)$  is easily seen to be

$$\phi(\omega) = -N \cdot 2\pi \log_e \omega + K \quad (10)$$

where  $K$  is a constant of integration. A plot of  $\phi(\omega)$  is shown in Figure 3. If we assume that all frequencies from  $\omega_L$  to  $\omega_U$  are not attenuated and all frequencies outside that band are totally attenuated the desired transfer function is as follows:

$$H(j\omega) = |H(j\omega)| \exp(j\phi(\omega)) = \exp(-jN \cdot 2\pi \log_e \omega + K) \quad (11)$$

for  $\omega_L \leq \omega \leq \omega_u$

$H(j\omega) = 0$  elsewhere

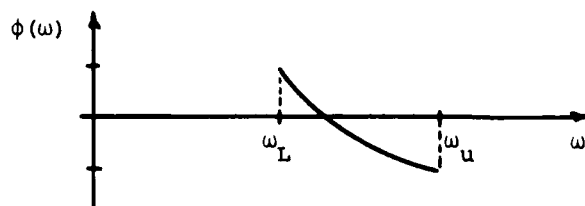


Figure 3 Plot of the phase of a logarithmically modulated phase signal.

If we further assume a real system then  $H(j\omega)$  for negative frequencies is the complex conjugate of that given in equation (11)

The impulse response of the system can then be found by taking the inverse Fourier transform of equation (11).

$$h(t) = \mathcal{F}^{-1}[H(j\omega)] = \int_{-\infty}^{\infty} H(j\omega) \exp(j\omega t) d\omega / 2\pi \quad (12)$$

Substituting  $H(j\omega)$  from equation (11) we see

$$h(t) = \text{Real} \int_{\omega_L}^{\omega_u} \exp(-jN \cdot 2\pi \log_e \omega + jK + j\omega t) d\omega / 2\pi \quad (13)$$

$$h(t) = \int_{\omega_L}^{\omega_u} \cos(N \cdot 2\pi \log_e \omega - K - \omega t) d\omega / \pi \quad (14)$$

A closed form solution for  $h(t)$  was not found. One approach for finding  $h(t)$  is to approximate the integral by using the principle of stationary phase which leads to an answer in terms of Fresnel integrals. These integrals must be evaluated by table or

numerically determined. Another approach is to use the inverse discrete Fourier Transform of the sampled function. A more direct approach is to numerically evaluate the integral for each  $t$  using Simpson's rule or Gauss quadrature.

#### B. Piecewise Logarithmic Modulated Phase Signal

The logarithmic modulated phase signal causes a delay inversely proportional to frequency; however, to be useful as an extrapolater the steady state output should be in phase with the input. Therefore the phase must be adjusted to zero or a multiple of  $2\pi$ . The group delay and phase requirements compete and we are not ever able to satisfy both requirements unless the phase is zero. One approach is to keep the approximate group delay the same as before and modify the phase by adding a stair case function. This is equivalent to considering the phase to be piecewise logarithmic. Such a phase is shown in Figure 4.

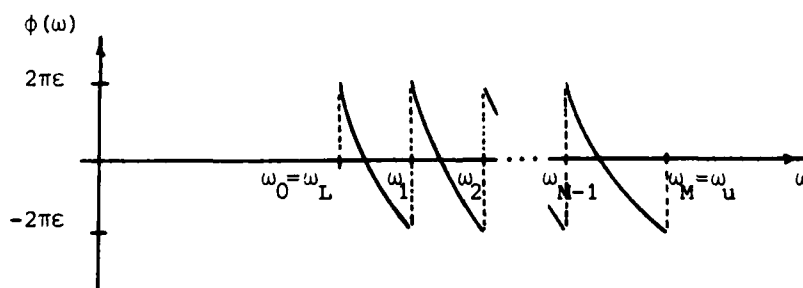


Figure 4 Phase characteristic for Piecewise logarithmically modulated Phase signal.

The band from  $\omega_L$  to  $\omega_u$  is broken into sections such that the

phase stays between  $\epsilon 2\pi$  and  $-\epsilon 2\pi$  where  $\epsilon$  is a factor specifying the maximum desired steady state phase error. The number of sections  $M$ , the corresponding  $\omega_i$  and  $K_i$  can be found for a specified  $\epsilon$ . Evaluating  $\phi(\omega)$  from Equation (10) at radian frequencies  $\omega_i$  and  $\omega_{i+1}$  we have.

$$\phi(\omega_i) = -N2\pi \log(\omega_i) + K_i = \epsilon 2\pi \quad (15)$$

$$\phi(\omega_{i+1}) = -N2\pi \log(\omega_{i+1}) + K_i = \epsilon 2\pi \quad (16)$$

By adding equation (15) and equation (16),  $K_i$  can be shown to be

$$K_i = N\pi \log(\omega_{i+1} \omega_i) \quad (17)$$

By subtracting equation (16) from equation (15),  $\omega_{i+1}$  can be found in terms of  $\epsilon$  to be as follows:

$$\omega_{i+1} = \omega_i \exp(2\epsilon/N) \quad (18)$$

For a given lower radian frequency  $\omega_L$  and upper frequency  $\omega_u$ , the number of segments can be shown to be.

$$M = [(N \log(\omega_u / \omega_L)) / 2\epsilon] \quad (19)$$

where  $[.]$  means next higher integer.

### C. Quadratic Phase

Another system which could provide longer extrapolation times is the one for which the delay is linearly proportional to the frequency as follows:

$$\tau_\omega = \alpha\omega \quad (20)$$

The slope of the phase is then given by

$$-d\phi(\omega)/d\omega = \alpha\omega \quad (21)$$

Integrating  $\phi(\omega)$ , equation (21) gives

$$\phi(\omega) = -\alpha\omega^2/2 + K \quad (22)$$

Therefore assuming constant magnitude throughout the bandwidth

the transfer function  $H(j\omega)$  becomes

$$H(j\omega) = \begin{cases} \exp(-j\alpha\omega^2/2 + K) & \omega_L \leq \omega \leq \omega_u \\ \exp(j\alpha\omega^2/2 - K) & -\omega_u \leq \omega \leq -\omega_L \\ 0 & \text{elsewhere} \end{cases} \quad (23)$$

The impulse response can then be obtained by integration see reference [8] for details. If we require that the steady state phase be within plus and minus  $\epsilon 2\pi$  then we must use a piecewise quadratic phase.

#### D. Piecewise quadratic phase

To obtain a given bandwidth,  $\omega_u$  and  $\omega_L$ , and satisfy the approximate group delay and steady state phase characteristic the phase must be sectionalized. The phase appears similar to that of Figure 4 except the sections are quadratic instead of logarithmic. At the endpoints of each interval we must have

$$\phi(\omega_{i+1}) = -\alpha\omega_{i+1}^2/2 + K = \epsilon 2\pi \quad (24)$$

$$\phi(\omega_i) = -\alpha\omega_i^2/2 + K = -\epsilon 2\pi \quad (25)$$

Adding the two equations the constant  $K$  is easily found to be:

$$K = \alpha(\omega_{i+1}^2 + \omega_i^2)/2 \quad (26)$$

Subtracting the two equations  $\omega_{i+1}$  can be shown to be

$$\omega_{i+1} = (\omega_i^2 - 8\epsilon\pi)^{1/2} \quad (27)$$

The number of segments  $M$  can then be found by letting  $\omega_0 = \omega_L$  and using equation (27) iteratively until  $\omega_i$  is greater than  $\omega_u$ .

## V. RESULTS

A computer program was written to simulate the systems with two of the impulse responses described in the previous section. Those chosen were the logarithmic phase and piecewise logarithmic phase signals. To determine the responses, the convolution integral was evaluated using Simpsons rule rather than using the Fast Fourier Transform.

Typical inputs and outputs are shown in figure 5 for specific parameters of the logarithmic phase impulse response. The parameter  $N$  of equation (11) could be adjusted resulting in an approximate delay for a sinusoidal signal input of exactly  $N$  cycles of whatever the input frequency is. However, the phase of the steady state output with respect to the input varies depending upon the frequency of the input sinusoid as expected by examining the phase of the system function. If  $\omega_u$  and  $\omega_L$  are not widely separated, it is possible to keep the steady state phase within  $\pm 0.2\pi$  as previously shown. If  $N$  is large, this becomes impossible for a fixed  $\omega_u$  and  $\omega_L$ .

It was found that the delay of the stepped sinusoidal input for the piecewise logarithmic impulse response did not show a delay of  $N$  times the period of the input signal. Typical results are shown in figure 6. By forcing the steady state phase to be within a given tolerance, we have lost the ability for a variable delay and experience only a small variable delay. Thus, the extrapolation length is limited to a small number of periods of the highest frequency considered which does not provide useful extrapolation.



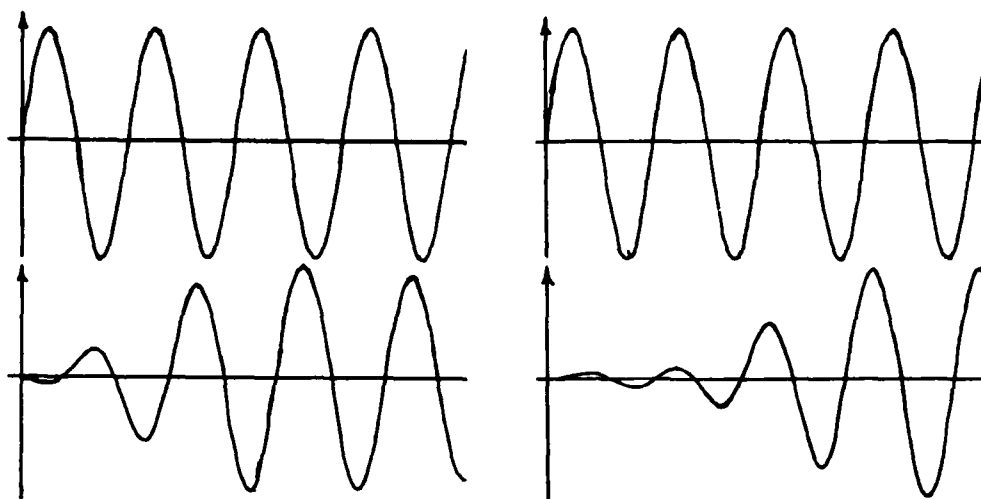


Figure 5. Plots of input  $x(t)$  and output  $y(t)$  for logarithmic phase transfer function given by equation (11) with  $f_L = 1$  Hz,  $f_u = 2$  Hz, input frequency 1.5 Hz, and (a)  $N = 1$ ,  $K = 17.83091$ ; (b)  $N = 2$ ,  $K = 29.37863$

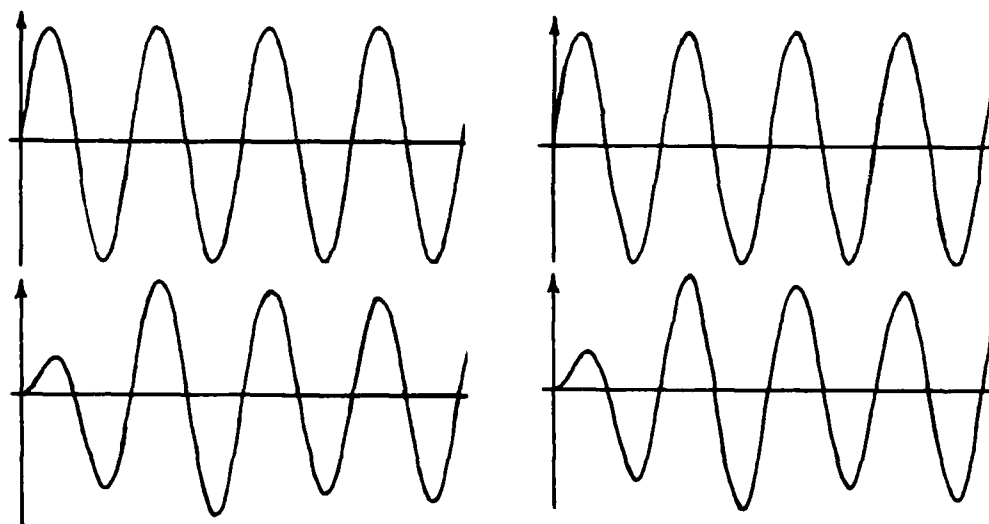


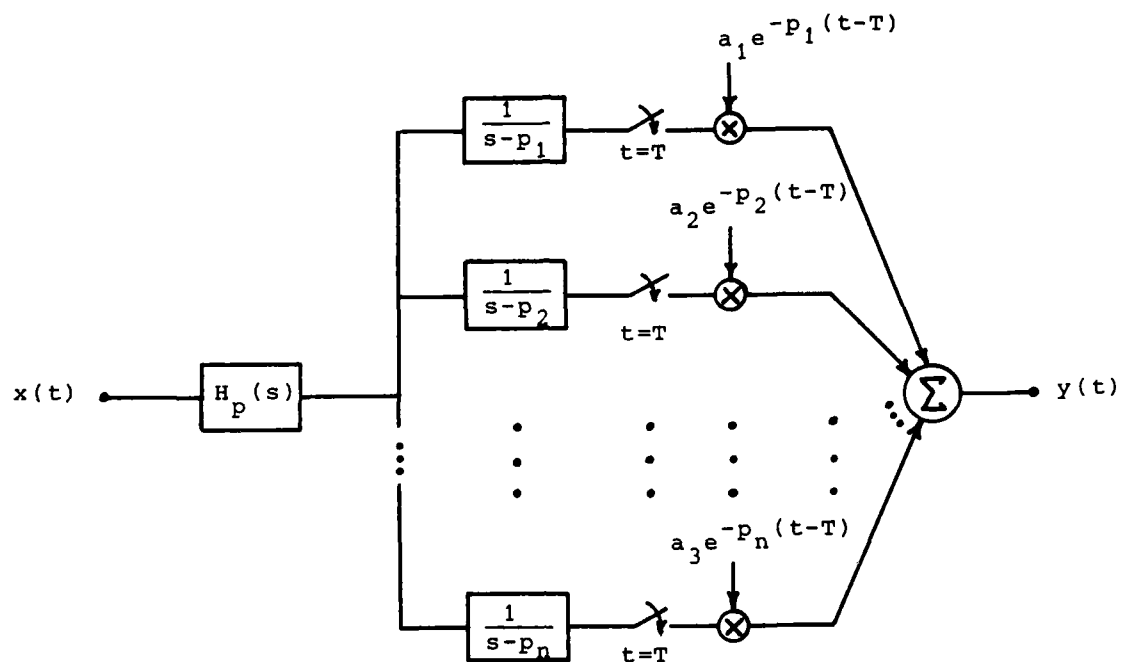
Figure 6. Plots of input  $x(t)$  and output  $y(t)$  for piecewise logarithmic transfer function specified by equations (15)–(18) with  $f_L = 1$  Hz,  $f_u = 2$  Hz, input frequency = 1.5 Hz,  $\epsilon = .1$ , and (a)  $N = 1$ , 4 segments; (b)  $N = 2$ , 7 segments.

## VI. RECOMMENDATIONS

It was shown that only limited extrapolation properties could be obtained by the systems examined. However, other suboptimum filters (with respect to the Wiener and Kalman filter) may be shown to provide useful extrapolation properties. It is recommended that the following approaches be investigated.

One approach would be to modify the logarithmic modulated phase signal by relaxing the causality condition and adding a linear phase characteristic to produce a specific gross time delay for extrapolation purposes.

Another approach would be to examine the structure shown in figure 7 which uses a prefilter followed by a bank of weighted low pass filters with specific cut-off frequencies. The values of these frequencies might be logarithmically space or suggested by the eigen functions of the Karhunen-Loève expansion for band-limited or low-pass random signals. An alternative structure would be a bank of weighted band-pass filters.



**Figure 7.** Possible extrapolator motivated by the Wiener filter for extrapolation.

#### REFERENCES

1. Hildebrand, F. B., Extrapolation, Interpolation, and Stationary Time Series, New York: Wiley, 1950.
2. Burg, J. P., "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Dept. of Geophysics, Stanford University, CA, May 1975.
3. Papoulis, A., "A New Algorithm in Spectral Analysis and Bands Limited Extrapolation", IEEE Trans on Circuits and Systems, Vol CAS-22, #9, pp 735-742, Sept. 1975.
4. Bowling, S. B., "Linear Prediction and Maximum Entropy Spectral Analysis for Radar Applications", Project Report RMP-122, MIT Lincoln Laboratory, May 24, 1977.
5. Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series, New York: Wiley, 1950.
6. Meditch, Stochastic Optimal Linear Estimation and Control, New York: McGraw-Hill, 1969.
7. Thor, T. C., "A large Time-Band width Product Pulse-Compression Technique, IRE Transactions on Military Electronics, Vol MIL-6, April 1962, pp 169-173.
8. Kroszczyiski, J. J., "Pulse Compression by Means of Linear Period Modulation, Proc of the IEEE, Vol 57, No. 7, July 1969, pp. 1260-1266.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

SILANE-TREATED SILICA FILLERS FOR ELASTOMER REINFORCEMENT

Prepared by:	Larry M. Ludwick
Academic Rank:	Professor
Department and University:	Department of Chemistry Tuskegee Institute
Research Location:	AFWAL/MLBT Wright-Patterson AFB
USAF Research:	Tom L. Graham
Date:	August 20, 1984
Contract No:	F49620-82-C-0035

## SILANE-TREATED SILICA FILLERS FOR ELASTOMER REINFORCEMENT

by

Larry M. Ludwick

### ABSTRACT

Fumed silicas, Cab-O-Sil M5 and EH5, were surface treated with mixtures of hexamethyldisilazane (HMDS) and 1,3-divinyltetramethyldisilazane (DVMS). Variations were made in the amount of silazanes used, the ratio of the vinyl and methyl silazanes, in the method of dispersing the filler during treatment and in the surface area of the filler. All treated fillers were then compounded with Silastic LS420. Following cure the formulations were subjected to standard tests to determine the effectiveness of the treatment. The high surface area filler EH5 treated with a 95 mol% HMDS/5 mol% DVMS solution involving a toluene reflux produced the greatest reinforcement effect in the elastomers. Using a colloid mill a large batch of the M5 filler was prepared for future use in the development of a new fluorosilicone hybrid elastomer known as FASIL.

#### ACKNOWLEDGEMENT

The author would like to acknowledge the support of the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for the opportunity to work on this project in the Materials Laboratory of the Air Force Wright Aeronautical Laboratories. The summer was made most enjoyable and beneficial by the congenial and stimulating atmosphere provided by the members of the Elastomer Group of the laboratory. Jerry Sieron, Warren Griffin, William Warner and particularly Tom Graham, the Effort Focal Point, provided much needed collaboration and guidance for this work. The support provided by the University of Dayton Research Institute (UDRI) is also appreciated. Tom Wical willingly contributed advice and was instrumental in designing the recirculating system for the colloid mill. The teaching and assistance of Wayne Polley made possible the preparation and testing of the many elastomeric compounds required for the evaluation of the candidate treated silica fillers.

## I. INTRODUCTION

A formulated silicone rubber consists of the polysiloxane polymer network, small particulate filler materials, additives for special colors or properties and a vulcanizing agent to provide chemical crosslinks. The filler components serve two basic functions. Certain fillers provide primarily an increase in bulk and thereby can serve to reduce cost. Other fillers play a more active role in the elastomer formulation and provide reinforcement and enhanced physical properties. A wide variety of both types of fillers is readily available. [1,2,3] An almost endless variety of elastomer formulations can be produced to match the physical properties of the rubber to the necessary specifications of an application.

In a filled elastomer, in addition to the polymer-polymer crosslinks and the effective crosslinks due to entanglement, the presence of a filler introduces the possibility of filler-polymer crosslinks, filler-filler crosslinks, polymer-filler weak interactions and polymer-filler entanglements. Of the reinforcing fillers, synthetic silicas are the most important for silicone rubbers. These very small particle silicas are of two types, fumed and precipitated, based on their method of preparation. The reinforcement effect arises from the structure of the filler aggregate itself, the particle size or surface area available, and the extent to which the polymer backbone and filler surface area interact, physically and chemically. [4,5,6] The surface of silica fillers is relatively well-defined, consisting of hydrophobic and hydrophilic areas. The hydrophobic areas comprise nonpolar silicon-oxygen-silicon, or siloxane, linkages. Silanol groups, either



isolated or adjacent (and thus hydrogen-bonded), result in hydrophilic regions. The reactive silanol groups result in ready association with the polysiloxane chains and can produce a premature hardening or crepe aging of a formulation before during. The hydrophilic silanol groups also increase the difficulty in dispersing the filler in the polysiloxane elastomer.

Numerous studies have shown that chemical modification of the silica filler surface can result in greatly increased reinforcement in elastomers. [7,8,9,10] Treatment of the silanol groups with various organosilanes, particularly trimethylsilane, shows reduced crepe aging. In addition, without the polar silanol groups, larger quantities of the fillers are more readily dispersed in elastomers. With the choice of appropriate organosilanes it is also possible to introduce built in crosslinking sites to further enhance physical properties. Thus the establishment of a properly modified silica filler plays a central role in the formulation of an elastomer possessing optimum physical properties.

## II. OBJECTIVES

The purpose of this project is to establish a treated filler with optimum properties for use in the development of a new fluorosilicone high temperature elastomer. A fumed silica, primarily Cab-O-Sil M5, will be treated with active disilazanes to produce a chemically modified filler. The quantity of the disilazane will be varied, according to the surface area of the filler, to theoretically provide partially or completely treated surfaces. The effect of the filler surface area will be investigated by comparing the results obtained with silica fillers having greater surface area, such as Cab-O-Sil EH5, with those of the M5

filler. Vinyl silanes will be incorporated in various percentages to determine the effect of the additional crosslinking sites. The filler will be dispersed in the solvent by different processes to determine the effect on the surface area available for treatment. The treated fillers will be compounded with a fluorosilicone polymer and the physical properties of the vulcanized materials evaluated.

### III. EXPERIMENTAL

#### MATERIALS

Polymers - Silastic LS420 (Dow Corning Corp.) a polysiloxane with methyl, trifluoropropyl and vinyl groups attached. PNF 200 (Firestone Tire and Rubber Co.) a phosphonitrilic fluoroelastomer.

Fillers - Cab-O-Sil M5 and Cab-O-Sil EH5 (Cabot Corporation). QUSO WR 82 (PQ Corporation).

Treating Agents - Hexamethyldisilazane (HMDS) and 1,3-Divinyltetramethyldisilazane (DVMS) were used as obtained from Petrarch Systems, Inc.

Compounding Ingredients - Luperco CST (50% 2,4-dichlorobenzoyl peroxide in silicone oil used with the fluorosilicone elastomer), Vul Cup 40KE (used with the PNF elastomer), zinc 8-hydroxyquinolate and Elastomag 170 were all used as obtained from suppliers.

Solvents - Toluene and methylene chloride were used as obtained from various suppliers. Both solvents were occasionally recycled by distillation.

#### FILLER TREATMENT

The filler was suspended in the minimum amount of toluene. Typical amounts would be 25 g of filler suspended in about 700 ml of solvent.

Water was added (5 parts per hundred parts filler) and the solution mixed thoroughly. Three methods were used to suspend the filler in the solvent: a mechanical paddle stirrer, a Waring Blender and a Charlotte Colloid Mill. The colloid mill was modified so that the mixture could be recirculated through a mixing funnel.

The silazanes were added as a solution with the solvent. Blending continued for several minutes to assure complete mixing. The viscosity of the suspension decreased so markedly that splashing in the blender container became a problem and blending could be maintained for only a few minutes after the addition. The solutions were then left for a period of time (usually overnight) before filtration. The suspensions were filtered using a Buchner funnel. The solid residue was air dried and then given a final drying for 4 hours at 150°C before use. Methylene chloride served as solvent for several trials. In these trials the solid was often collected by evaporating the methylene chloride.

Modifications in the filler treatment procedure included trials where the vinyl silane was added to the suspension in pure form. The vinyl silane was added first with the addition of the trimethyl silane following several hours later. A second modification involved heating the reaction mixture to the reflux point of toluene. At approximately 85°C the silica toluene mixture became clear as though the silica were dissolving.

#### COMPOUNDING

All compounding was done by hand mixing the ingredients on a 3" x 8" two-roll mill. Recipes for the two elastomer formulations and the curing conditions are given below.

<u>1</u>	Silastic LS420	100
	silica filler	30
	Lupercu CST	2
	Press Cure: 5 minutes at 240°F	
	Post Cure: 16 hours at 400°F	

<u>2</u>	PNF 200	100
	Silica Filler	30
	Elastomag 170	2
	Zinc 8-Hydroxyquinolate	1.5
	Vul Cup 40 KE	0.6
	Press Cure: 30 minutes at 340°F	
	Post Cure: 4 hours at 350°F	

#### PHYSICAL PROPERTIES

The following properties were measured on cured samples of each elastomer. The silastic compounds were molded as 0.050 inch sheets. Tear specimen, dumbbells and buttons for the compression set measurement were die cut from these sheets. The PNF compounds were molded as O-rings for all test except tear samples which were cut from 0.050 inch sheets.

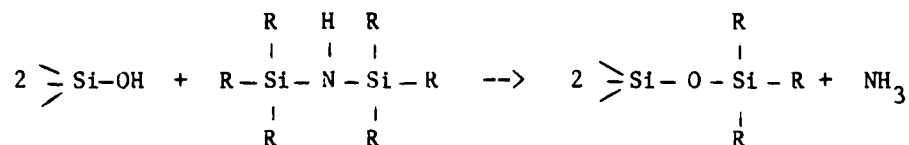
Hardness  
Tear Strength, Die B  
Tensile Strength, 1/8 inch (constricted width) dumbbell  
Elongation  
50% Modulus  
100% Modulus  
Compression Set (25% compression, 72 hours at 275°F)

For many samples the tensile strength and elongation were also measured at 275°F using an Instron Tensile Tester with an oven mounted on one frame. Elongation required opening the oven door during the test. However, the temperature of the sample would not have decreased significantly during the time required for the test.

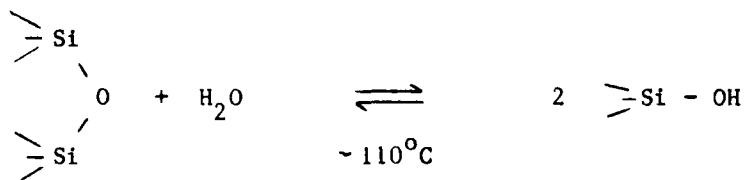
#### IV. RESULTS

The chemical reaction between the silanol group of the filler and

the disilazane occurs according to the following scheme.



With a silanol concentration of  $4.5 \text{ OH/nm}^2$  a treatment level of 0.75 mol disilazane/kg filler was used. This is sufficient to assure complete reaction of the available silanol groups. All reactions were carried out in the presence of additional water to promote the conversion of siloxane groups to free silanols.



With the application of the disilazanes to the filler several obvious changes occurred. The viscosity of the mixture was rapidly reduced, to the point that severe splashing resulted in the blender container. When the filler was dried there was a reduction in the volume of the solid. Although no exact measurements were made, the volume occupied by the 15 g sample used in formulations was reduced to approximately 25% of the volume of the untreated filler. Also in the compounding process the treated fillers dispersed much more readily in the elastomers, producing products that were very tacky, frequently sticking to the rolls of the mill. Based on this observation it appeared that a much larger amount of the treated filler could have been compounded with the elastomers. The untreated fillers produced stiff

products that could be readily removed from the roll and then cut to shape with scissors. Finally, the treated fillers were hydrophobic as opposed to the easy wettability by water of the untreated fillers.

As shown in Table I, treatment of Cab-O-Sil M5 with a mixture of disilazanes (95 mol% hexamethyldisilazane and 5 mol% 1,3-divinyldisilazane) leads to a significant increase in the tensile strength and percent elongation in the Silastic LS420. In addition the modulus, hardness and compression set show significant decreases. At 275°F the changes in tensile strength and elongation are maintained, although each is reduced by approximately 50%. It appears that the use of a mechanical stirrer, a blender or a colloid mill to disperse the filler for treatment does not significantly alter the properties of the compounded elastomer. There are however some practical differences in that the colloid mill is capable of processing larger quantities of the filler than is possible with the other two methods.

Table II shows the same set of treated fillers compounded with the phosphonitrilic fluoroelastomer (PNF 200). Cab-O-Sil M5 does not represent a good filler for PNF 200 as noted by the compression set values of nearly 100%. Even though the elongation and tensile properties are higher for the untreated fillers, the material showed no return from the 25% deformation at 275°F after 72 hours. With the treated fillers, the compounds are softer and more elastic but do not exhibit greater tensile strength. Because of the constraints of time and material, no other compounding was done with PNF 200.

The results obtained with filler treated with two different DVMS concentrations are shown in Table III. In each case the quantity of silazanes used to treat the filler was varied systematically. Two fillers treated without the addition of water to the mixture are also

TABLE I

Silane Treated Cab-O-Sil M5 (Compounded with Silastic LS420)

Compound #	Si-1438	-1439	-1440	-1441	-1443	-1444
Mixing Process <sup>1</sup>	MS	MS	B	B	CM	CM
Silane Treatment <sup>2</sup> mol silane/kg filler	0	0.75	0	0.75	0	0.75
<u>Property</u>						
Hardness, Shore A	71	55	71	57	71	56
Tear (lb/in)	87	97	83	97	87	107
Tensile Strength (lb/in <sup>2</sup> )	960	1280	910	1250	980	1250
Elongation (%)	--- <sup>3</sup>	260	--- <sup>3</sup>	260	190	280
Modulus, 50% (lb/in <sup>2</sup> )	295	170	324	204	361	193
100% (lb/in <sup>2</sup> )	497	347	580	437	604	378
<u>at 275°F</u>						
Tensile Strength (lb/in <sup>2</sup> )	476	611	441	680	497	560
Elongation (%)	90	170	80	179	100	140
25% Compression Set, 72 hr (%)	30	17	27	18	31	12

<sup>1</sup>MS = Mechanical Stirrer, B = Waring Blender, CM = Colloid Mill

<sup>2</sup>95 mol% HMds, 5 mol% DVMS

<sup>3</sup>Less than 100%, actual value not measured

TABLE II

Silane Treated Cab-O-Sil M5 (Compounded with PNF 200)

Compound #	PN-1448	-1449	-1451	-1452	-1453	-1454
Mixing Process <sup>1</sup>	MS	MS	B	B	CM	CM
Silane Treatment <sup>2</sup> mol silane/kg filler	0	0.75	0	0.75	0	0.75
<u>Property</u>						
Hardness, Shore A	80	50	80	57	80	52
Tear (lb/in)	245	153	215	191	253	168
Tensile Strength (lb/in <sup>2</sup> )	1390	1160	1550	1400	1430	1120
Elongation (%)	300	210	300	220	240	200
Modulus, 50% (lb/in <sup>2</sup> )	389	129	340	184	400	184
100% (lb/in <sup>2</sup> )	467	340	464	363	600	378
<u>at 275°F</u>						
25% Compression Set, 72 hr (%)	98	56	100	64	100	54

<sup>1</sup>MS = Mechanical Stirrer, B = Waring Blender, CM = Colloid Mill  
<sup>2</sup>95 mol% HMDS, 5 mol% DVMS



TABLE III  
Variation in Treatment Levels of Cab-O-Sil M5 (Compounded in Silastic LS420)

Compound #	Si-1448	-1449	-1450	-1463	-1452	-1453	-1454	-1464
<u>Silane Treatment</u> <sup>1</sup>								
mol% DVMS	5	5	5	5 <sup>2</sup>	10	10	10	10 <sup>2</sup>
mol silane/kg filler	0.38	0.75	1.50	0.75	0.38	0.75	1.50	0.75
<u>Property</u>								
Hardness, Shore A	61	58	56	60	61	56	56	61
Tear (lb/in)	97	90	92	88	104	88	99	85
Tensile Strength (lb/in <sup>2</sup> )	1140	1160	1010	982	941	916	959	885
Elongation (%)	200	220	210	220	210	200	230	210
Modulus, 50% (lb/in <sup>2</sup> )	227	222	203	204	251	228	200	240
100% (lb/in <sup>2</sup> )	418	448	405	381	458	423	353	430
at 275°F								
25% Compression Set (%)	17	18	12	18	18	18	17	21
<sup>1</sup> Blender treatment								
<sup>2</sup> No water added during filler treatment								

included in Table III. The results of all trials appear very similar with the exception of a significant decrease in tensile strength for those fillers treated with a silazane mixture containing 10% of the vinyl compound.

Table IV shows the results of adding the vinyl silazane separately. Again the differences appear slight. However, there is a suggestion of a decrease in tensile strength with increasing vinyl content. In the case of these trials the silazanes were added in neat form rather than as a dilute solution. This addition method may result in a less uniform distribution of the silane groups on the filler surface even though the additions were made with the blender running. The addition of a solution of premixed silazanes appears to be the more successful treatment method.

The effect of surface area on reinforcing properties is shown in Table V with the use of Cab-O-Sil EH5. This filler has nearly double the surface area ( $390 \pm 40 \text{ m}^2/\text{g}$ ) as compared to Cab-O-Sil M5 ( $200 \pm 25 \text{ m}^2/\text{g}$ ) used in all previous trials. The tensile strength and elongation increase markedly with increased filler treatment. Although tear strength appears to decrease with increased treatment, all values are greatly elevated over those of the untreated filler. In addition, the moduli and compression set show significant decreases with silazane treatment. It may also be significant that the best values for all properties, with the exception of tear and high temperature tensile strength, occur for the sample that had been heated to toluene reflux temperature for several hours. Near reflux temperatures the filler appears to dissolve due to the removal of air and more complete wetting of the surface by the solvent. Although this treatment was inadvertant and only intended to remove solvent by distillation, it appears that the more complete dispersion of the filler resulted in a more uniformly

TABLE IV

Vinyl Silane Treated Cab-O-Sil M5 (Compounded in Silastic LS420)

Compound #	Sil-1460	-1461	-1462	-1463	-1464
Silane Treatment <sup>1</sup>					
mol silane/kg filler	0	0.75	0.75	0.75	0.75
% DVMS <sup>2</sup>	0	0	1	5	10
<u>Property</u>					
Hardness, Shore A	71	59	61	60	61
Tear (lb/in)	86	86	85	88	85
Tensile Strength (lb/in <sup>2</sup> )	522	1110	901	982	885
Elongation (%)	100	200	180	220	210
Modulus, 50% (lb/in <sup>2</sup> )	311	205	218	204	240
100% (lb/in <sup>2</sup> )	550	428	448	381	430
<u>at 275°F</u>					
Tensile Strength (lb/in <sup>2</sup> )	314	427	528	482	506
Elongation (%)	60	110	120	120	120
25% Compression Set, 72 hr (%)	28	19	17	18	21

<sup>1</sup>Blender treated

<sup>2</sup>DVMS added directly, then HMDS to complete treatment level

TABLE V

Silane Treated Cab-O-Sil EH5 (Compounded in Silastic LS420)

Compound #	Sil-1456	-1457	-1458	-1459	-1465
Silane Treatment <sup>1</sup> mol silane/kg filler	0	0 <sup>2</sup>	0.9	1.9	2.0 <sup>3</sup>
<u>Property</u>					
Hardness, Shore A	73	71	58	54	56
Tear (lb/in)	86	82	127	120	118
Tensile Strength (lb/in <sup>2</sup> )	496	675	1200	1330	1540
Elongation (%)	100	160	310	310	320
Modulus, 50% (lb/in <sup>2</sup> )	353	308	193	176	147
100% (lb/in <sup>2</sup> )	495	484	359	329	285
<u>at 275°F</u>					
Tensile Strength (lb/in <sup>2</sup> )	284	452	485	767	672
Elongation (%)	50	110	130	200	290
25% Compression Set, 72 hr (%)	39	45	16	16	11

<sup>1</sup>95 mol% HMDS, 5 mol% DVMS, blender treated<sup>2</sup>Dried at 150 C before use<sup>3</sup>Cab-O-Sil EH5 heated to toluene reflux during treatment

treated product. Preliminary tests indicate a similar result occurs with the M5 silica filler as well.

V. RECOMMENDATIONS:

This work has demonstrated significant increases in the reinforcing properties of fumed silica fillers in a fluorosilicone elastomer by surface treating the filler material with silanes. However, a number of questions remain as to the exact effect of the chemical treatment and its relationship to the observed physical properties of the compounded elastomer.

The silane treatment of a filler with larger surface area showed greater reinforcement (Cab-O-Sil EA5 versus Cab-O-Sil M5). This should be firmly established with additional preparations. Other fillers with a variety of surface area characteristics should be treated to establish the optimum surface area for reinforcing the elastomer properties.

Although a vinyl silane component was used to treat the fillers in order to provide crosslinking sites, the actual amount incorporated was not established. The number of vinyl pendant groups on the fillers should be experimentally established and related to the properties of the cured elastomers. It would be desirable to establish exactly how much of the silanes were actually incorporated. Thus far the amount incorporated has been assumed based on a quantitative reaction of the starting material used in each treatment procedure.

Other aspects of the treatment procedure need further investigation, primarily the effect of heating the reaction mixture with the resulting increased wetting of the filler surface. Also the optimum time for the reaction may be different from the "overnight" period frequently used.

Another extension of this study would look at the structure of the

treated and untreated fillers, both alone and as compounded. Electron microscopy could be used to determine whether or not the treating process has had any significant effect on the aggregate structure. The use of torn and stretched samples of the cured elastomers would also yield information about the surface alterations in this treating process.

#### REFERENCES

1. "Compounding with Silastic<sup>(R)</sup> Gums and Basic," Dow-Corning Publication, 1968.
2. E. M. Dannenberg, "Filler Choices in the Rubber Industry," Rubber Chem. Technol., 55 , 861 (1982).
3. M. P. Wagner, "Reinforcing Silicas and Silicates," Rubber Chem. Technol., 49 , 703 (1976).
4. E. M. Dannenberg, "The Effects of Surface Chemical Interaction on the Properties of Filler-Reinforced Rubber," Rubber Chem. Technol., 48 , 410 (1975).
5. K. E. Polmanteer and C. W. Lentz, "Reinforcement Studies-Effect of Silica Structure on Properties and Crosslink Density," Rubber Chem Technol., 48 , 410 (1975).
6. B. B. Boonstra, H. Cochrane and E. M. Dannenberg, Rubber Chem Technol, 48 , 558 (1975).
7. E.P. Plueddemann, "Silane Coupling Agents," Plenum Press, New York, 1982.
8. E. M. Dannenberg and G. R. Cotten, Rev. Gen. Gaoutch, Plast. 51 , 347 (1974).
9. M. T. Mason and C. L. Lee, "Effects of Fumed Silica Treated with Functional Disilazanes of Silicone Elastomers Properties," Rubber Chem. Technol., 55 , 233 (1982).
10. E. L. Warrick, O. R. Pierce, K. E. Polmanteer and J. C. Saam, "Silicone Elastomer Development 1967-1977," Rubber Chem. Technol., 52 , 437 (1979).

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE APPLICATION OF STRUCTURAL EQUATION MODELING TO

EXPERIMENTAL DATA

Prepared by:	Robert MacCallum
Academic Rank:	Associate Professor
Department and University:	Psychology Department Ohio State University
Research Location:	Air Force Aerospace Medical Research Laboratory, Human Engineering Division, Technology Development Branch
USAF Research:	Sharon Ward
Date:	August 22, 1984
Contract No:	F49620-82-0035



THE APPLICATION OF STRUCTURAL EQUATION MODELING  
TO EXPERIMENTAL DATA

by

Robert MacCallum

ABSTRACT

The potential application of a quantitative procedure called structural equation modeling (SEM) to experimental data is investigated. Conditions under which such a use of SEM could be of value are specified. The use of coded variables to represent experimentally manipulated independent variables is described, and approaches to the incorporation of such variables into structural equation models are proposed. A method for adjusting data from repeated-measures designs, so as to make the data more suitable for analysis via SEM, is suggested. A brief illustration is presented. General recommendations are offered, along with suggestions for further research.

#### ACKNOWLEDGMENTS

I would like to express appreciation to the Air Force Systems Command, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for the opportunity to spend a very productive summer conducting research at the Aerospace Medical Research Laboratory at Wright-Patterson AFB, Ohio. I would like to thank the personnel of the Human Engineering division, particularly the Technology Development branch, for providing a friendly and stimulating environment. Especially helpful and supportive were Sharon Ward, Walt Summers, Maris Vikmanis, Cliff Brown, and John Forester.

## I. INTRODUCTION

The author's background lies in the field of quantitative methods, particularly methods of research design, measurement, and data analysis in the behavioral sciences. A very important development in this field during the past ten years is represented by a technique called structural equation modeling (SEM) (Bentler, 1980; Long, 1983). SEM is a technique which provides a capability for fitting and testing theories about patterns of relationships among variables. In a typical application, a researcher hypothesizes that certain measured variables (MV's) are actually approximate measures of latent variables (LV's, hypothetical constructs which cannot be directly measured), and that there exists a particular pattern of effects among these LV's. Given a sample of observations on the MV's, SEM then allows the investigator to estimate the strength of all of the effects and relationships in the hypothesized model, and to assess whether the model is a plausible explanation of the data.

Researchers at AFAMRL/HEC have recently become interested in the potential value of SEM in their research. In much of their research, a hypothesized prior model involving LV's exists. However, they felt that their data analysis procedures did not provide an adequate theory testing capability nor an explicit means for incorporating LV's.

However, a major stumbling block in their use of SEM is that SEM was designed for and is generally used in the analysis of non-experimental data, whereas the research at AFAMRL/HEC is almost exclusively experimental. That is, their research involves experimental manipulation of independent variables, and subsequent observation of dependent variables. SEM, on the other hand, is intended for fitting and testing models using non-experimental data, in which both independent and dependent

variables are sampled. Thus, the very general goal of the author during the research period was to investigate the feasibility of applying SEM to experimental data.

## II. OBJECTIVES OF THE RESEARCH EFFORT

This general goal included a series of specific objectives:

- 1) To first evaluate on a logical and conceptual basis the conditions under which the application of SEM to experimental data would be of any potential value. Given that such data are almost universally treated via analysis of variance (ANOVA), it was important to first determine whether and when SEM could provide any advantage over ANOVA.
- 2) To develop a technique for incorporating experimentally manipulated variables into SEM. As will be seen below, a variety of issues must be considered in the use of such categorical variables in SEM.
- 3) To develop a technique for treating data from repeated measures experimental designs in SEM. Since virtually all of the experimental research at AFAMRL/HEC employs repeated-measures designs (i.e., observation of the same subject under different experimental conditions), and since SEM assumes observations to be independent, some procedure needs to be employed to "correct" the data for the presence of the effects of the repeated-measures factor(s) in the design.
- 4) To carry out applications of SEM to some data collected at AFAMRL/HEC, for the purpose of illustrating the general SEM procedure as well as the specific procedures for analyzing experimental data. In addition, this would serve to demonstrate the use of the LISREL computer program (Joreskog & Sorbom, 1981), a powerful program for fitting and testing structural equation models.

### III. FEASIBILITY AND VALUE OF APPLYING SEM TO EXPERIMENTAL DATA

A pre-requisite for the application of SEM to experimental data to be of any interest is that there be more than one measured dependent variable. Given this, there are at least two general conditions under which a model-testing approach to experimental data could be useful. The first would be the case in which the investigator wished to test some path-analytic model involving relationships among the experimentally manipulated independent variables and the measured dependent variables. The second would be the case in which one or more subsets of the measured dependent variables were intended to serve as multiple indicators of latent variables. Of most interest and value would be the situation in which these two conditions were combined, resulting in a model postulating a specific pattern of relationships among experimentally manipulated independent variables and latent dependent variables, where there are multiple indicators for each of the latter.

The general point here is that in order for anything to be gained via the application of SEM to experimental data, there must exist some over-identified model to be tested. If there is only one measured dependent variable, or if there is more than one but no testable structural model or measurement model, then a confirmatory analysis via SEM is inappropriate. Univariate and/or multivariate ANOVA will generally suffice. If, however, a testable model is present, then the use of SEM offers substantial advantages over the more traditional ANOVA approach. Foremost among these is the use of a confirmatory, rather than exploratory, approach to the data analysis. SEM provides a direct, explicit, and integrative investigation of the hypothesized model, along with a rich array of information about goodness of fit and potential model modifications. ANOVA offers no such

capability or information. Other clear advantages of SEM over ANOVA (when a testable model is available) include (a) the capability for representing latent variables, and (b) the explicit representation of error of measurement in measured variables. These aspects of SEM can be contrasted with ANOVA, in which (a) there is no capacity for directly representing latent variables, and (b) all measurements are treated as error-free, an assumption which can lead to bias in results (Bentler, 1980).

#### IV. REPRESENTING EXPERIMENTALLY MANIPULATED VARIABLES IN SEM

A problem in the application of SEM to experimental data is the matter of how to incorporate the experimentally manipulated independent variables into the data and the model. It is relevant here to distinguish between two types of such variables: (a) those whose levels correspond to different numerical values of some quantitative variable (e.g., drug dosage); and (b) those whose levels correspond to different categories with no numerical meaning (e.g., drug A vs. drug B). Variables of the first type can be incorporated very easily into SEM data and models. With respect to the data, such a factor can be represented by a single measured variable, with each subject's "score" on the variable given by the numerical value associated with the given level of the factor. In a model, this measured variable will be hypothesized to have a particular pattern of effects on other variables. The incorporation of the second type of experimentally manipulated variables (i.e., those which are truly categorical) into SEM is more complex. With respect to the data, the mechanism for representing categorical independent variables is the same as that commonly used to represent such variables in regression analyses; i.e., the construction of coded variables. The actual definition and construction of coded variables will not be discussed in detail here.

Standard texts on regression analysis routinely cover this topic; especially good sources are Cohen and Cohen (1983, pp. 181-222) and Pedhazur (1982, pp. 269-333).

Both of these texts discuss three different methods for constructing coded variables: dummy coding, effects coding, and contrast coding. In the present context, the choice of coding methods will be a function of whether the investigator's hypotheses include specific statements about the exact nature of the effect of the categorical independent variable on the other variable(s) in the model. For instance, one might hypothesize that the difference between level 1 and level 3 of a categorical independent variable affects a particular dependent variable. If the investigator's model includes such hypotheses, then contrast coding should be used so that coded variables can be constructed to represent the particular contrasts of interest. If, instead, the model hypothesizes overall effects of a categorical variable on other variables, without being specific about the exact nature of those effects, then either dummy coding or effects coding would be appropriate.

Regardless of which coding method is used, for a categorical independent variable with  $J$  levels, one must construct  $J-1$  coded independent variables (e.g., dummy variables). If a study includes more than one categorical independent variable, such a set of coded variables must be constructed for each categorical independent variable. It is not necessary to use the same coding method for each categorical variable. The resulting complete data matrix would then include all the necessary coded variables along with all of the measured non-manipulated variables. The latter category will typically correspond to all of the dependent variables. However, to be completely general, this category could

additionally include some measured variables which are not experimentally manipulated but which serve as independent variables (or as indicators of latent independent variables). In any case, a covariance matrix for all of the coded and measured variables can be obtained and can serve as input for a SEM analysis.

The incorporation of the coded variables into a model must be approached carefully. There seems to be a variety of potential logical and conceptual trouble-spots. For instance, when an experimentally manipulated variable is represented by a set of dummy variables, it may be tempting to consider the dummy variables to be multiple indicators of a latent variable representing the experimental manipulation. This representation is illustrated in Figure 1, where three coded variables (C1, C2, and C3 being coded variables for a four-level factor) are shown as multiple indicators of a single independent latent variable. However, such a representation can be seen to be inappropriate when one considers the mathematical framework of SEM. Suppose, for this example, that the coded variables were created via dummy coding. According to the equations for the measurement model, as shown in Figure 1, such a model would imply that each dummy variable is equal to some unknown weight multiplied by the latent variable. Clearly, it is impossible to multiply a single latent variable by different weights to produce a set of dummy variables.

Rather than consider coded variables to be multiple indicators of latent variables, it seems more appropriate to employ coded variables directly as measured variables in a model. That is, each coded variable would serve as a separate independent variable and, based on the investigator's model, would exhibit a particular pattern of effects on the dependent variable(s) in the model.



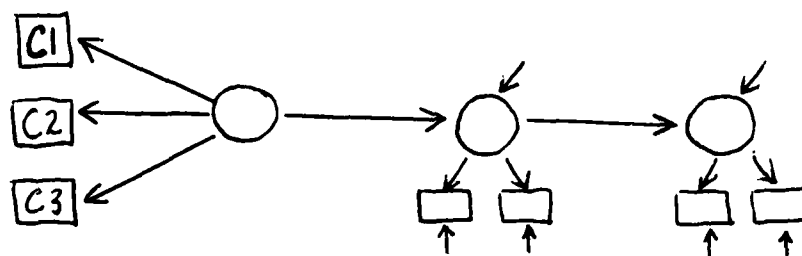


Figure 1

An Incorrect Use of Coded Variables:  
Multiple Indicators of a Latent Variable

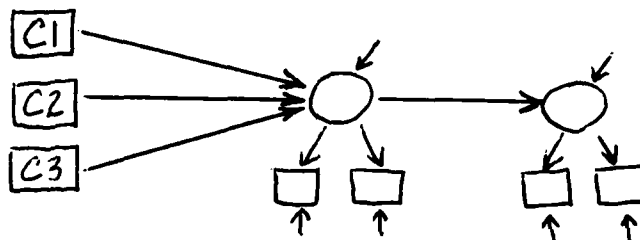


Figure 2

A Model Showing Same Effects  
for all Coded Variables Within a Set

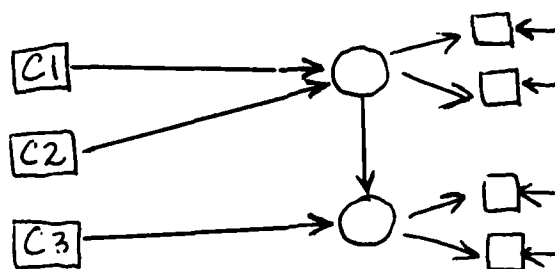


Figure 3

A Model Showing Different Effects  
for Coded Variables Within a Set

It is useful to differentiate here between two possible strategies involving defining a model of the effects of the coded variables on the dependent variable(s). One strategy would be appropriate when a model includes an hypothesis that a given categorical independent variable affects a given dependent variable, without the specific nature of the effect being postulated. Such a model would then include an effect of each of the coded variables (for that categorical variable) on the given dependent variable. An example of such a model, using three coded variables to represent a four-level categorical variable, is shown in Figure 2. In a model of this type, there would be a path coefficient associated with the effect of each coded variable on the dependent latent variable. Interpretation of such coefficients would be analogous to the interpretation of regression weights for coded variables, as discussed by Cohen and Cohen (1983) and Pedhazur (1982).

The critical point here is that, in the absence of specific hypotheses about the nature of the effect in question, the coded variables representing the given factor are treated as a set in the initial model. The solution for the initial model might suggest subsequent modifications involving the addition or deletion of effects for only a subset of these coded variables. This is permissible, but the investigator in such a case should carefully consider what each coded variable represents, so that the modified model and associated parameter estimates can be interpreted correctly.

A different general strategy for incorporating coded variables into an initial model would be useful when that model included specific hypotheses about the nature of the effect of the categorical factor on the dependent variable(s). Given contrast coded variables to represent the specific

effects of interest, the initial model would then include effects for the coded variables so as to correspond to the investigator's specific hypotheses. In such a model, the coded variables representing a given factor would not necessarily be treated as a set. A model of this type is illustrated in Figure 3, in which it is hypothesized that one dependent LV is affected by two of the contrasts, while a different dependent LV is affected by a third contrast. This approach offers much greater flexibility and finer control and interpretation of the modeling process than does the previous strategy. However, its use is dependent on the investigator having the necessary well-justified a priori hypotheses.

#### V. ADJUSTING DATA FROM REPEATED-MEASURES DESIGNS

The basic data to which a SEM analysis is applied is the covariance matrix among all the MV's. When this matrix is calculated using all the observations from a repeated-measures design, the covariances will be based on multiple observations from each subject. Such covariances will undoubtedly be affected by between-subject effects, and it would be desirable to remove such effects from the data prior to analysis by SEM. The author proposes that an appropriate procedure for accomplishing this would be to partial the subject effects from each dependent variable, and to employ the resulting residuals (the "corrected" dependent variables) in the calculation of the desired covariances. Following is a description of a specific procedure for carrying out this process. Users should keep in mind that this needs to be done only when there is one or more within-subjects factor in a design.

The key to the entire process is to define a set of coded variables to carry the information about the arrangement of subjects in the experimental design. Each subject measured under each condition (or combination of

conditions) will be assigned a value, or "score," on each of these coded variables. The coded variables can be defined via the following steps:

1) Do steps (2) and (3) once for each level, or combination of levels, of the between-subjects factors. Each repetition of steps (2) and (3) defines a subset of coded variables. If there is no between-subjects factor, steps (2) and (3) are carried out just once. If there is only one between-subjects factor, these steps are repeated once for each level of that factor. If there is more than one between-subjects factor, these steps are repeated once for each cell involving the between-subjects factors, ignoring the within-subjects factor(s).

2) Let  $N_j$  be the number of subjects within the level or cell being considered.

3) Define  $N_j - 1$  coded variables using "effects coding" as follows, where  $v_{ik}$  is the value of coded variable  $k$  for the  $i$ 'th subject in the level or cell being considered:

a) for the first coded variable,

let  $v_{i1} = 1$  for  $i = 1$   
                   $= -1$  for  $i = N_j$   
                   $= 0$  for all other  $i$

b) for the second coded variable,

let  $v_{i2} = 1$  for  $i = 2$   
                   $= -1$  for  $i = N_j$   
                   $= 0$  for all other  $i$

c) generally, for the  $k$ 'th coded variable,

let  $v_{ik} = 1$  for  $i=k$   
                   $= -1$  for  $i = N_j$   
                   $= 0$  for all other  $i$

If these steps are repeated as defined in (1), then the total number of coded variables will equal the total number of subjects, minus the number of levels or cells of the between-subjects factor(s).

Once the coded variables representing subject effects are defined as above, they are then used to partial out subject effects from the dependent variables via a series of regression analyses. These analyses can be accomplished as follows:

- 1) Given  $r$  coded subject vectors  $X_1, X_2, \dots, X_r$ .
- 2) Given  $t$  dependent variables  $Y_1, Y_2, \dots, Y_t$ .
- 3) Carry out  $t$  separate multiple regression analyses, one for each of the  $t$  dependent variables. In each of these analyses, a given dependent variable  $Y_s$  will be regressed on the  $r$  coded subject variables, which are treated as independent variables. Thus, each of these regression analyses will yield a variable  $Y^*_s$ , which is the value of  $Y_s$  predicted from a linear combination of the coded subject variables.
- 4) For each of the  $t$  dependent variables, obtain residuals by subtracting predicted values from observed values. Let

$$e_s = Y_s - Y^*_s$$

The "residualized dependent variable"  $e_s$  will then represent that portion of  $Y_s$  which remains after subject effects have been removed. There will be  $t$  such residualized variables  $e_1, \dots, e_t$ .

- 5) These residualized dependent variables can then be used in place of the original measured dependent variables in the computation of the covariance matrix for the entire set of MV's. All resulting covariances involving dependent variables will have been corrected for the presence of subject effects. This covariance matrix can then be used as input to a SEM analysis.

## VI. APPLICATION TO REAL DATA

Using the procedures described in this report, SEM was applied fairly successfully to data collected from two different experiments at AFAMRL/HEC. Results from one of these studies will be described briefly. In a study of the effects of jamming and deception on information processing and decision making in subjects playing the role of an Air Defense Commander, it was hypothesized that the various dependent measures obtained were indicators of two distinct LV's which could be called uncertainty and performance. The former represents the indecision or confusion exhibited by a subject as he makes decisions, and the latter represents the subject's skill and success in carrying out the required task. The initial model stated that the experimental manipulation of jamming and deception would affect uncertainty, which would in turn affect performance. Since the data were from a repeated-measures design, the subjects' scores on the dependent variables were corrected via the approach described in Section V above. In addition, coded variables were defined to represent the jamming/deception variable, as described in Section IV. The model stated above was then fit to the resulting data via the LISREL computer program. Results indicated that, while uncertainty and performance seemed to be reliable, valid LV's, the model relating these LV's to the jamming/deception manipulation fit the data very poorly. Commonly used procedures for model modification (Joreskog & Sorbom, 1981) were employed in order to determine whether a modified version of this model would fit the data better, and a model which fit the data quite well was found. This model indicated that performance was affected by uncertainty and by the manipulation of the jamming/deception conditions. Interestingly, the model implied that uncertainty was not affected by

jamming and deception, but rather that it was operating as an uncontrolled independent variable.

## VII. RECOMMENDATIONS

The author would strongly encourage the use of SEM for the analysis of experimental data. The techniques described above provide a workable approach to such analyses, as illustrated in the previous section. It is clear that much experimental research, such as that conducted at AFAMRL/HEC, is characterized by the presence of a prior model involving LV's, and that SEM offers distinct advantages over the traditional ANOVA approach in such cases, as discussed in Section III above. More specifically, results of such a modeling procedure may be of great value in military-related research, where it may be desirable to determine how to optimize or manipulate individuals' performance or response. This objective becomes more reachable with a clear understanding of the "causal network" which relates the relevant variables.

The author would also urge potential users of SEM to view this technique with an emphasis on its confirmatory nature. That is, it is most useful to consider SEM as an integrative research technique, beginning with the formulation of a model and proceeding through research design, data collection, and data analysis.

It is clear, however, that SEM can also be used in an exploratory fashion. Indeed, it can be quite informative to apply SEM to data which may not have been collected for that purpose, and/or to make modifications in an initial model so as to enhance the parsimony or goodness of fit of the model. In fact, it became clear during the author's stay at AFAMRL/HEC that the greatest value of SEM to researchers there would lie in such an approach, since most of their research was more exploratory than

confirmatory in nature. Such a use of SEM can clearly be beneficial with respect to identifying LV's and developing or refining roughly defined models.

However, it should be recognized that such an approach gives rise to a potential for capitalization on chance. If the formulation of the initial model were based substantially on post hoc, rather than a priori information, or if model modifications were excessive in number or lacking in substantive meaning, the end product of the analysis might be largely a function of chance characteristics of the data. The investigator can gain some protection against this problem by (a) attempting to provide strong a priori justification for the initial model; (b) limiting model modifications to those which both are substantively meaningful and bring about a significant improvement in the parsimony and/or fit of the model; and (c) cross-validating the final model(s) on a second sample or in a follow-up study (Bentler, 1980). In exploratory applications of SEM, all of these guidelines are quite important to follow and should enhance the validity of resulting models and conclusions.

The process of model modification is an especially critical issue here. When an initial model is found to fit poorly, it is common practice, as mentioned in Section VI, to attempt to modify the model so as to improve its fit. An exploratory application of SEM may involve a series of such modifications. This raises an interesting and important question: If there exists some "true" model different from the researcher's initial model, will the commonly used modification procedures necessarily lead to the correct model? Unfortunately, the author has produced some preliminary information suggesting that the answer to this question may be "no." The author is submitting an application for a Research Initiation grant to



support a detailed investigation of this issue, including an examination of the performance of various strategies for model modification. Results of this research should be very relevant to researchers using SEM in an exploratory manner.

### References

- Bentler, P. M. "Multivariate analysis with latent variables." Annual Review of Psychology, 1980, pp. 419-456.
- Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum, 1983.
- Joreskog, K. G., & Sorbom, D. LISREL User's Guide. Chicago: International Educational Services, 1981.
- Long, J. S. Covariance structure models: An introduction to LISREL. Beverly Hills: Sage Publications, 1983.
- Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, 1982.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

THE USE OF THE INSTANTANEOUS FREQUENCY TRANSIENT IN THE DESIGN AND  
OPTIMIZATION OF THE CHANNELIZED RECEIVER AND INSTANTANEOUS  
FREQUENCY MEASUREMENT (IFM) VERSIONS OF THE PASSIVE EW RECEIVER

Prepared by:	Dr. William S. McCormick
Academic Rank:	Associate Professor
Department and University:	School of Engineering Wright State University
Research Location: Beam	Air Force Avionics Laboratory, Electronic Support Measures Group
USAF Research	Dr. James Tsui
Date:	August 17, 1984
Contract No:	F49620-82-C-0035

THE USE OF THE INSTANTANEOUS FREQUENCY TRANSIENT IN THE DESIGN AND  
OPTIMIZATION OF THE CHANNELIZED RECEIVER AND INSTANTANEOUS  
FREQUENCY MEASUREMENT (IFM) VERSIONS OF THE PASSIVE EW RECEIVER

by

William S. McCormick

ABSTRACT

Using analytic signal representation, the instantaneous frequency transient,  $W_I(t)$ , was analyzed for the filter/limiter chain of the High Probability Intercept Receiver (HPIR) in both the channelized receiver and IFM configurations. Using the concept of ensemble processing, the  $W_I(t)$  transient was used with a conventional FM discriminator to define a new processing scheme that holds the promise of operating at very low pulse widths. The concept of ensemble processing was also applied to envelope processing within the channelized receiver configuration using the amplitude histogram, the matched filter, and the laquerre expansion respectively. The frequency transient was also used to analyze and calculate the estimation bias of the IFM receiver in both the traditional and in the channelized receiver hybrid configurations.

AD-A154 337

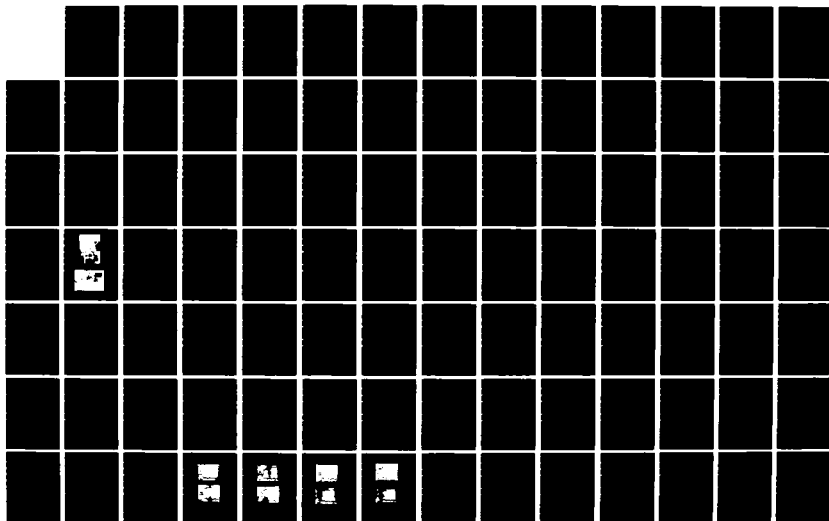
UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 12/13  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

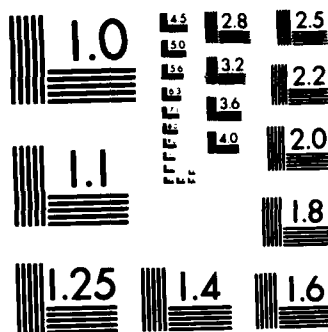
UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

#### ACKNOWLEDGEMENTS

The author would like to thank the Air Force Wright Aeroantualical Laboratories, the Air Force Office of Scientific Research, and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend an interesting and useful summer at the Avionics Laboratory, WPAFB, Dayton, Ohio. He would like to acknowledge the Electronic Support Measure Group for the excellent working condition provided and also for their generosity in providing whatever support was asked for.

The author would also like to give special thanks to Dr. J. Tsui who suggested the research topic and continued to lend his guidance and support throughout the project. Final thanks to John Sloan of Wright State University for his invaluable assistance during the computer programming phase.

## I. INTRODUCTION

The frequency estimation of enemy radar pulses is of vital and obvious importance to the U.S. Air Force. The passive electronic warfare (EW) device that performs this function is called the "High Probability Intercept Receiver (HPIR)".<sup>[1]</sup> The HPIR receiver is currently configured in either the Channelized Receiver (CR) mode or the Instantaneous Frequency Measurement (IFM) mode with both configurations possessing some relative advantages and disadvantages.

Briefly described, the channelized receiver (see Figures 1 to 3) is a heterodyne device that shifts a wide range of input frequencies (2.6 to 3.9 GHz.) into a range of frequencies covered by a contiguous bank of "slot filters" with bandwidths or frequency resolutions in the order of 20 MHz. The actual frequency estimation is made by monitoring the activity levels of the various "slot" filters using a variety of schemes that operate exclusively on the filter response envelopes. A major complication of the filter bank approach is the transient response of more than one slot to the leading edge of the input pulse. The principal function of the envelope processing circuitry is to retain the fundamental frequency resolution in the presence of this multiple slot transient response. Unfortunately, the exclusive reliance on the slot filter envelope response has resulted in a pulse width operating limit that is significantly greater than the theoretical lower bound.

The second configuration, the IFM receiver, estimates frequency by measuring at a given sampling time the differential phase,  $w_1 \tau$ , accumulated by passing the signal pulse through a fixed known time delay,  $\tau$ ; Figure 26 illustrates. The IFM receiver can operate at low pulse widths but it is extremely sensitive to any perturbation in the instantaneous frequency or phase of the input signal. Phase perturbations can result either from the instantaneous frequency transient or from the effects of multiple signals. It is clear that the proper analysis and optimization of both HPIR configurations requires a flexible simulation and analysis of the effects of the frequency transient and the effects of multiple signals. This need is given added urgency by the



current trend toward using the IFM as the fine frequency measuring device within a channelized receiver configuration.

Two major HPIR simulation programs have been written. The first HPIR simulation was performed at the Georgia<sup>[2]</sup> Tech Engineering Station over the period 1973 to 1982. While this simulation package simulated a broad range of HPIR processing steps, it was never debugged fully and it never addressed directly the effects of the filter frequency transient on the IFM/channelized receiver configuration. In addition, it was computationally inefficient since it performed the simulation at RF and did not use the more computationally efficient baseband, analytic signal, low-pass representation. The analytic signal, equivalent loss-pass representation was however used in a second major simulation effort performed by the University of Missouri and System Research laboratories over the period 1975 to 1984.<sup>[3]</sup> The deliverable item in this second project was an experimentally verified program written in BASIC for the HP 9286 desk computer that could handle a broad range of operations, including multiple signals, in a user-friendly interactive mode. Unfortunately, the instantaneous frequency response was never isolated and displayed in this program nor was any direct attention paid to the error effects on the HPIR receiver.

## II. OBJECTIVES

The main objective of this project was the analytic signal analysis of the HPIR receiver configuration with special attention given to the generation and system implications of the instantaneous frequency transients. In detailed, itemized form, the specific objectives are as follows:

1. To develop a complete analysis and simulation of a filter-limiter chain with special attention given to the instantaneous frequency transient. The simulation will include the multiple signal case and will be based on analytic signal analysis.
2. To investigate the interrelationship of the instantaneous frequency transient to the time variation of the in-phase and quadrature components.

3. To simulate the "Double Detection" scheme of the conventional channelized receiver and to evaluate its performance; to also consider new envelope-based frequency measurement schemes based on a more general parameter estimation approach.
4. To use the instantaneous frequency transient,  $\omega_I(t)$ , response to analyze the IFM receiver and CR/IFM hybrids.
5. To consider new methods of short pulse discrimination using the  $\omega_I(t)$  response.

### III. ANALYTICAL DISCUSSION

Taking the signal center frequency as a reference, the asymmetrical low-pass transfer function can be written, for an offset of  $\Delta$ , as

$$H(\omega - \Delta\omega) = H_p(\omega - \Delta\omega) + j H_q(\omega - \Delta\omega) \quad (1)$$

where  $H(\omega)$  is the filter loss-pass equivalent transfer function and  $H_p(\omega - \Delta\omega)$ ,  $H_q(\omega - \Delta\omega)$  are the low-pass transfer functions of the in-phase and quadrature components of the impulse response. [4]

The  $h_p, h_q$  components of the asymmetrical low-pass impulse response can be determined directly by recognizing the Fourier transform pair given as follows,

$$H(\omega - \Delta\omega) \leftrightarrow e^{-j\Delta\omega t} h(t) = h_p(t) + j h_q(t) \quad (2)$$

where  $h(t)$  is the low-pass impulse response of the filter. Using Euler's identity, the asymmetrical  $h(t)$  reduces to

$$\begin{aligned} h_p(t) &= h(t) \cos(\Delta\omega t) \\ h_q(t) &= -h(t) \sin(\Delta\omega t) \end{aligned} \quad (3)$$

It is noted that  $h_q(t) = 0$  when  $\Delta\omega = 0$  as expected. The operation defined in (3) formed the basis for the asymmetrical transient analysis.

In general, a signal input with in-phase and quadrature components,  $P_s(t)$ ,  $Q_s(t)$ , applied to a filter with complex impulse response ( $h_p(t)$ ,  $h_q(t)$ ) will [5] give in-phase and quadrature output components,  $y_p(t)$ ,  $y_q(t)$ , given, with \*

representing convolution, by

$$\begin{aligned} (y_p(t), y_q(t)) &= (P_s, q_s) * (h_p, h_q) \\ &= ([P_s * h_p - q_s * h_q], [P_s * h_q + q_s * h_p]) \end{aligned} \quad (4)$$

or

$$\begin{aligned} y_p(t) &= P_s * h_p - q_s * h_q \\ y_q(t) &= P_s * h_q + q_s * h_p \end{aligned}$$

The filter output,  $y(t)$ , referenced to the signal center frequency,  $\omega_0$ , can be written as

$$y(t) = 2 y_p(t) \cos(\omega_0 t) - 2 y_q(t) \sin(\omega_0 t) \quad (5)$$

which reduces, by phasor addition, to

$$y(t) = 2 \sqrt{y_p^2 + y_q^2} \cos[\omega_0 t + \tan^{-1} \{ \frac{y_q}{y_p} \}] \quad (6)$$

where the envelope,  $e(t)$ , is given as

$$e(t) = 2 \sqrt{y_p^2 + y_q^2} \quad (7)$$

and the phase angle,  $\theta(t)$ , as

$$\theta(t) = \tan^{-1} \left[ \frac{y_q}{y_p} \right] \quad (8)$$

The instantaneous frequency,  $\omega_I(t)$ , can be obtained by differentiating (8) to get<sup>[4]</sup>

$$\omega_I(t) = \frac{d\theta}{dt} = \frac{\frac{y_p \dot{y}_q - y_q \dot{y}_p}{y_p^2}}{\frac{y_q^2}{y_p^2 + y_q^2} + 1} = \frac{y_p \dot{y}_q - y_q \dot{y}_p}{y_p^2 + y_q^2} \quad (9)$$

Expression (9) is a key expression in the work reported here.

An interesting observation relating envelope response and instantaneous frequency can be made from expression (9). Defining  $\hat{A}$

as  $(y_p, y_q)$  and  $\vec{B}$  as  $(\dot{y}_p, \dot{y}_q)$ , the vector cross product of  $\vec{A}$  and  $\vec{B}$  can be expressed as

$$\vec{A} \times \vec{B} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ y_p & y_q & 0 \\ \dot{y}_p & \dot{y}_q & 0 \end{vmatrix} = \underline{k}(y_q \dot{y}_p - y_p \dot{y}_q) \quad (10)$$

=  $\underline{k}|\vec{A}||\vec{B}|\sin \gamma$  where  $\gamma$  is the angle between  $\vec{A}$  and  $\vec{B}$ .

Recognizing  $|\vec{A} \times \vec{B}|$  as the numerator of expression (9) for  $\omega_I(t)$ , we can write

$$\omega_I(t) = \frac{|\vec{A}||\vec{B}|\sin \gamma}{y_p^2 + y_q^2} \quad (11)$$

Defining  $E(t)$  as the square of the envelope, we have

$$E(t) = y_p^2 + y_q^2 \quad ; \quad (12)$$

forming the derivative yields

$$\dot{E}(t) = 2y_p \dot{y}_p + 2y_q \dot{y}_q \quad (13)$$

which can be interpreted (Dot Product) as

$$\dot{E}(t) = 2\vec{A} \cdot \vec{B} = 2|\vec{A}||\vec{B}|\cos \gamma$$

which then reduces expression (9) to a new and useful relationship,

$$\omega_I(t) = \frac{1}{2} \frac{\dot{E}(t)}{E(t)} \tan(\gamma) \quad (14)$$

Expression (14) relates envelope and instantaneous frequency response and emphasizes the importance of the relative orthogonality of  $(y_p, y_q)$  and  $(\dot{y}_p, \dot{y}_q)$ . In particular, it points out that a significant frequency transient may exist even in a relatively "quiet" region of envelope response ( $\dot{E}(t)=0$ ) provided  $\vec{A}$  and  $\vec{B}$  are nearly at right angles ( $\tan \gamma = \infty$ ). Since the attainment of the steady state forces  $\vec{A}$  and  $\vec{B}$  to be orthogonal ( $\dot{E} = 0 = y_p \dot{y}_p + y_q \dot{y}_q$ ), this interpretation may explain the violent frequency transient observed at the trailing edge of the filter output since the state of orthogonality must decay over the same time constants as the filter envelope.

#### IV. Simulation Details and Program Verification

All filters are Chebyshev with the low-pass impulse response,  $h(t)$ , generated by taking the inverse FFT of the transfer function as defined by the pole positions. In discussing the signal flow, a filter-limiter-filter cascade is represented as  $(F_1 L F_2)$  with  $(BW_1/BW_2)$  representing the respective RF bandwidth (MHz) of the filters. The time resolution of the simulation is 2.5 nanoseconds.

The limiter stage is considered as a zero-memory device with the instantaneous frequency of the output equaling the instantaneous frequency of the input. This condition is imposed by maintaining the same ratio between in-phase and quadrature terms at the input and output. This assumption is equivalent to neglecting the higher harmonics which is quite valid for a narrow-band system. The limiting is performed on the input envelope using the ARCTAN function defined as follows:

$$\text{output envelope} = (2/\pi) \tan^{-1}[k(\text{Normalized Input Envelope})] \quad (15)$$

where  $k$  is adjusted according to the "hardness" of the limiting action. In practice, a 50% condition is maintained where a voltage 50% of peak value attains saturation.

Consistent with the model defined by equations 1 thru 15, a flexible FORTRAN program was written to provide envelope, instantaneous frequency, and the orthogonality measure,  $\gamma$ , for a general filter-limiter cascade. The model used was low-pass in nature with the zero frequency representing the center frequency of the input RF signal pulse. By varying the offset,  $\Delta\omega$ , in an appropriate manner, the simulation was able to handle heterodyning and filtering at any level in the overall system. In order to verify the program, a comparison was made with the experimentally verified REC BASIC program which unfortunately only provided envelope response. For a typical and trapezoidal signal of 300 n.s. duration with 30 n.s. leading and trailing edges, the response of a  $F_1 L F_2$  chain (20 MHz/40 MHz) is given in figures 4 thru 12 for an offset of 16 MHz.

The significance of Figures 4 thru 12 are explained in the following items:

1. Figure 4 presents a comparison of the envelope of the first filter ( $F_1$ ) of bandwidth 20 MHz to the output of the REC (Inset) program. Excellent agreement is obtained which lends confidence to the accuracy of the  $y_p(t)$ ,  $y_q(t)$  components. The unmarked time legend for the REC plot is 400 n.s. per division. The vertical scale of the figure is of no special significance and is only retained to identify limiter action.
2. Figure 5 presents the instantaneous  $\omega_I(t)$  (equation 9) at the output of filter 1. The frequency is defined relative to a signal frequency that is 16 MHz greater than the filter center frequency. Because the leading edge of the input pulse appears as a delta function to the filter, the initial  $\omega_I(t)$  will be near the center frequency of the filter; it will then approach zero as the  $\omega_I(t)$  response approaches zero signal frequency in the steady state. The transients in the envelope in the offset case produce characteristic "ears" as is observed in Figure 4. Returning to Figure 5, the  $\omega_I(t)$ , as expected, goes negative and then approaches zero until the violent transient at  $t = 150$  n.s. This transient is explained from equation (14) where the "dip" in envelope causes a large  $\dot{E}(t)$  value and also from the orthogonality measure,  $\gamma$ , which goes through  $\pi/2$  at  $t = 150$  n.s. After the transient, the  $\omega_I(t)$  approaches its steady state value of zero. It is noted that the violent peaks at  $t = 150$  n.s. is the result of a phasing relationship between  $(y_p, y_q)$  and  $(\dot{y}_p, \dot{y}_q)$  and does not occur at lower offsets. Continuing, the  $\omega_I(t)$  undergoes a small peaking at the trailing edge [ $t = 300$  n.s.] and then drops to the value of -12 MHz and finally oscillates about -6 MHz. This interesting and potentially useful trailing edge transient is due to the  $\dot{E}(t)$  and to a  $\gamma$  value near  $\pi/2 = 90$  degrees.

3. Figure 6 presents the orthogonality measure,  $\gamma$ , for the first filter output. The interesting points are at the  $t = 150$  n.s. time and also during the trailing edge interval.
4. The "HARD" limited output of filter #1 is shown in Figure 7. It is noted that the major "dip" in envelope at  $t = 150$  n.s. has been flattened by the limiting action. The rise in side-lobe levels for  $t \geq 400$  n.s. is also noted. Good agreement with the REC plot is observed.
5. Figure 8 presents the  $\omega_I(t)$  plot for the limiter output. As expected, it is the same as the limiter input. However, since the envelope change,  $\dot{E}$ , is much smaller, the  $\gamma$  angle must reside near  $\pi/2$  to ensure the same  $\omega_I(t)$ . This is confirmed by the plot of Figure 9.
6. Figure 10 gives the envelope response of the second 40 MHz filter,  $F_1$ . The agreement with the REC inset is quite good. Figure 11 presents the  $\omega_I(t)$  output of the second filter. The early negative peak has been reduced to -12 MHz from -9 MHz indicating the effect of the limiter.

It should also be mentioned that a multiple signal capability has been included in the program. Each pulse is defined independently in terms of pulse width, offset, and time shift. Using equation (3), the in-phase and quadrature components are summed for the signal components and the simulation proceeds to the next step.

#### IV. Uses of Simulation

##### A. Simulation of Double Detection Scheme for the Channelized Receiver

The usual frequency measurement component of the channelized receiver (CR) is the double detection circuit of Figure 12 which divides the incoming slot signal into two parallel paths with different bandwidth filters. When there is no frequency offset, the signal output of the narrower bandwidth filter will exceed the wider bandwidth filter. When a significant offset is present, the output of the wider bandwidth filter will exceed the output of the narrow band filter.

Provided the sampling time is chosen correctly, the double detection circuit can detect and eliminate those slot filters that are responding in the "offset" mode. Figures 13 and 14 illustrate a zero and 24 MHz offset case for a 300 n.s. pulse using the configuration of Figure 12. Although the scheme is effective if the envelope comparison is made at pulse center, the  $\Delta f = 0$  slot will not be identified if the comparison is made on the leading edge of the response. An even more basic problem occurs at short pulse widths where the required "ear" structure is lost for non-zero offset case as shown in Figures 15 and 16 for the 100 n.s. case.

The double detection method is a crude method for resolving the filter transient problem. The method is envelope response based and makes no direct use of the frequency transient. Furthermore, the circuit bases its decision solely on what is occurring within a given slot and does not consider the responses in the entire ensemble of slots. The concept of ensemble processing is a new approach that has great promise. The next section will advance some sophisticated envelope-based discrimination approaches that make use of the ensemble concept.

B. Envelope Based Frequency Detection Using the Ensemble Concept

(i) The Histogram

Forming the histogram or the relative frequency approximation to the probability density function is a common operation in digital image processing and non-parametric statistics. For the HPIR, the proposed approach will form an amplitude histogram of the detected envelope of each slot and compare it, in a mean squared error sense, to the expected histogram of the zero offset case. For  $N$  slots, there are  $N$  real number across the ensemble with the minimum value determining the true frequency; Figure 17 illustrates. Figures 18 and 19 illustrate the dramatic difference in histogram structure for a 300 n.s. pulse in a  $F_1LF_2$  cascade with  $F_1 = 20$  MHz and  $F_2 = 40$  MHz, and an offset that varies from 0 to 40 MHz. As expected, the  $\Delta f = 40$  MHz case has a greater



concentration at lower magnitudes with its fine structure varying as well as its magnitude range.

A histogram is a well-defined method of defining the amplitude signature of a signal. A microprocessor can easily be programmed to form a histogram. Unfortunately, a real-time implementation of histogram generation is not possible for the 200 n.s. range given current microprocessor speeds.

(ii) The Matched Filter Approach

Since the zero-offset envelope waveform,  $e_0(t)$ , is known apriori, it is of interest to consider the cross-correlation function,  $y_k(t)$ , of  $e_0(t)$  with each of the slot envelope functions,  $e_k(t)$ ; that is,

$$y_k(t) = \int_0^t e_0(\tau)e_k(\tau)d\tau. \quad (16)$$

The correlation or matched filter approach is the optimum maximum likelihood detector in additive gaussian noise. In the present context, it is proposed that the ensemble of  $y_k(t)$  be compared as a function of time in order to estimate the zero offset slot. Using the configuration of the histogram test, the cross correlations of (9) were calculated for a range of offsets; the pulse widths were taken to be 100 and 300 n.s. with the sampling time,  $\bar{t}$ , taken at 600 n.s. The considerably longer pulse widths at the output of the cascade result from a combination of limiter action and filter memory. The results are given in the following Table:

$y_k(\bar{t})$ (Normalized)			
(100 n.s.)		(300 n.s.)	
$\Delta f_k(\text{MHz})$	$y_k(\bar{t})$	$\Delta f_k(\text{MHz})$	$y_k(\bar{t})$
0	1.0	0	1.0
28	0.84	16	0.84
32	0.68	20	0.58
36	0.58	36	0.39
40	0.53	40	0.29

Table 1: Cross Correlations Versus Offset  
( $\bar{t} = 600$  n.s.)

For both pulse widths, the  $y_k(\bar{t})$  stays essentially constant for offsets less than the 20 MHz filter bandwidth. For the longer pulse width of 300 n.s., the ensemble variation with offset is more dramatic which again illustrates the short pulse problem.

(iii) Laquerre Coefficient Expansion

Waveforms can be efficiently represented by a coefficient vector where the coefficients are coefficients of a well chosen set of orthogonal basis function. One popular class of orthogonal basis function,  $l_n(t)$ , are the Laquerre functions which are orthogonal in the interval  $(0, \infty)$  and have an exponential character. The Laplace transform of the Laquerre functions,  $L_l(s)$ , is given in reference 5 as

$$L_n(s) = \frac{(S-p)^n}{(S+p)^{n+1}} = L\{l_n(t)\} \quad (17)$$

with  $p$  representing a scaling constant. Expression (17) is immediately recognized as a cascade of real-time, single role, all-pass networks as shown in Figure 20. The intriguing property of the Laquerre networks is that the convolution integral relating filter input to filter output can be interpreted as the integral defining the

Laquerre coefficient; that is, the nth coefficient,  $a_n$ , can be expressed as

$$a_n = \int_0^t e(\tau) l_n(\tau) d\tau \quad (18)$$

which can be considered as the filter response at the various taps or output points defined in Figure 20. It follows that the entire waveform can be represented as a coefficient vector  $(a_0, a_1, a_2, \dots, a_n)$  which suggests that a Euclidian distance measure,  $D_k$ , can be defined for slot k relative to the zero-offset slot as follows

$$D_k = \frac{\left| \sum_{i=1}^n (a_i^0 - a_i^k) \right|}{\sqrt{\left\{ \sum_{i=1}^n (a_i^0)^2 \right\} \left\{ \sum_{i=1}^n (a_i^k)^2 \right\}}} \quad (19)$$

where  $a_i^k$  is the ith coefficient for the kth slot. Values for  $D_k$  are presented in Table II for various offsets for the 100 and 300 n.s. case using the previous configuration.

(100 n.s.)		(300 n.s.)	
$\Delta f_k$ (MHz)	D	$\Delta f_k$ (MHz)	D
0	0	0	0
8	5.05	8	0.34
16	1.70	16	6.63
24	16.45	24	42.45
32	70.36	32	99.77

Table II: Ensemble Variation of Laquerre Vectors

Except for the "dip" at  $\Delta f = 8$  MHz in the 100 n.s. case, the D variation with offset offers sensitive enough ensemble variation to suggest a potentially useful frequency measuring scheme.

The above candidate methods are merely examples of the ensemble approach. Many more could be suggested and will be investigated in future work.

C. Short Pulse Methods Using the Instantaneous Frequency Transient

i. Use of a Conventional FM Discriminator to External Operation to Short Pulse Widths

From the frequency responses  $\omega_I(t)$ , it was observed that significant frequency transient activity occurred at short pulse widths for no-zero offset. At these short pulse widths, the concurrent envelope responses are so distorted that a significant frequency estimation error occurs when standard envelope response methods (eg: double detection) are used. In the approach proposed, a nonsynchronous FM discriminator is used to estimate the frequency transient activity. The activity levels of the contiguous slot filter are then interpolated over the ensemble to estimate the correct signal frequency. The foundation for this method is the fact that there is no frequency transient activity whatsoever in the zero offset case (i.e., no quadrature component)!! It is proposed to use a conventional lead filter to convert the FM activity to envelope activity exactly as in a conventional FM discriminator. One possible FM discriminator would be a limiter at each slot followed by a simple lead network as shown in Figure 21. One undesirable quality of the above approach, however is that the lead network will closely approximate pure differentiation and that it will, as a result, generate an envelope derivative term due to the pulse nature of the input signal. This envelope derivative, which is not present in the commercial (stationary) FM applications, must be considered as a distortion term since it is only weakly coupled to offset value. In this scheme, it is proposed to measure this envelope derivative term and to then subtract it from the discriminator output thereby exposing the frequency transient activity. This approach will be made clear from the following analysis.

Consider the signal,

$$s(t) = E(t) \cos \{[\omega_I(t)]t\} \quad (20)$$

where  $E(t)$  is the envelope pulse response and  $\omega_I(t)$  is the instantaneous frequency transient which equals zero for zero offset. Forming the derivative of  $s(t)$ , we get using the chain rule,

$$\begin{aligned} \dot{s}(t) &= \overset{(A)}{\dot{E}(t)} \cos\{[\omega_I(t)]t\} \\ &\quad - \underset{(B)}{E(t)[\omega_I(t) + (\dot{\omega}_I(t))t] \sin\{[\omega_I(t)]t\}} \end{aligned} \quad (21)$$

where term (A) is identified as the envelope derivative term and term (B) is identified as the desired FM term. Since terms (A) and (B) are in quadrature, the envelope of  $s(t)$  can be written as follows,

$$\text{Env.}\{\dot{s}(t)\} = \sqrt{(\dot{E})^2 + E^2(t)(\omega_I(t) + t \dot{\omega}_I(t))^2} \quad (22)$$

It is proposed to isolate  $\dot{E}$  by first envelope detecting the output of the limiter in a parallel path and then passing that output through a rate filter equivalent to the filter described above. The isolated  $\dot{E}$  is then subtracted from the envelope of  $\frac{ds}{dt}$  as given below:

$$\left\{ \begin{array}{c} \text{output} \\ \text{of} \\ \text{subtractor} \end{array} \right\} = \left[ \sqrt{(\dot{E})^2 + E^2(t)(\omega_I(t) + t \dot{\omega}_I(t))^2} - \dot{E} \right]; \quad (23)$$

when  $\Delta f = 0$ ,  $\omega_I(t) = \dot{\omega}_I(t) = 0$ , and expression (23) reduces to  $\dot{E} - \dot{E} = 0$  as expected. The entire scheme is illustrated in the schematic of Figure 22.

The scheme has been programmed on the computer as an extension to the  $\omega_I(t)$  transient program. The lead networks were analyzed by convolving input signals with the

filter impulse response given  $h(t) = \delta(t) - 1/RC e^{-t/RC}$   
 $\mu(t)$ . Referring to the above, the integral of the absolute  
 value at point (F) is taken as the "activity index" which  
 increases across the ensemble with offset  $\Delta f$  as shown in the  
 computer results of Table III.

$T_p = 100 \text{ n.s.}$		$T_p = 300 \text{ n.s.}$	
$\Delta f(\text{MHz})$	I	$\Delta f$	I
0	0.0	0	0
4	0.143	4	1.69
8	0.473	8	4.81
12	0.879	12	7.48
16	1.282	16	9.23
20	1.625	20	6.97

Table III: Activity Levels for FM Discrimination Method.

The FM discriminator approach makes the first direct use of the frequency transient. By employing the ensemble concept and the idea of activity levels, the method presents an attractive approach to short pulse frequency detection.

ii. Direct Detection and Utilization of the In-Phase and Quadrature Components

The response of the filter bank to a given RF pulse can be considered as an ensemble of function pairs  $(y_p(t), y_q(t))$ . As demonstrated in digital communication theory, the optimum processor for the maximum likelihood estimator in white gaussian noise is the quadrature processor shown in Figure 23. This type of processor is commonly used in the QAM and QPSK modems and can be directly extended to the HPFIR application. Figures 24 and 25 present the integrals of the  $y_p$  and  $y_q$  components for the  $F_1LF_2$  (20/40) case with a pulse width of 300 n.s. for the zero offset and 24 MHz offset

case respectively. The dramatic difference between the two ( $y_p, y_q$ ) responses suggests a number of ways to define a useful activity index for ensemble processing.

A major assumption of the above computer runs was the assumption of phase coherence which is not an operationally practical assumption. A number of approaches could possibly resolve the phase coherence problem including the use of multiple references at uniform phase spacings. An investigation of this interesting approach will be proposed in the mini-grant proposal.

D. The Effect of the Instantaneous Frequency Transient on the IFM Receiver

i. General Remarks

As mentioned in the introduction, the IFM receiver measures at a given sampling time the differential phase,  $\omega_i \tau$ , that is accumulated when a signal is passed through a time delay,  $\tau$ . Through a sequence of mixing operations, the phase lag is isolated as  $A \sin(\omega_i \tau)$  and  $A \cos(\omega_i \tau)$  which leads directly to the frequency estimate  $\omega_i$  given by [3]

$$\omega_i = \left(\frac{1}{\tau}\right) \tan^{-1}[\tan(\omega_i \tau)]. \quad (24)$$

Figure 26 illustrates the general configuration while Figure 27 gives a more detailed description of the frequency measurement circuits. Accurate operation of the IFM receiver requires that the instantaneous frequency of the signal into the IFM device be constant and equal to the frequency of the input signal. Any phase perturbation due either to asymmetrical filter operation or the presence of multiple signals can lead to a phase perturbation or bias in the IFM output.

Referring to Figure 26, the undelayed and delayed can be written as follows:

$$\cos[\omega_0 t + \int_0^t \omega_I(\alpha) d\alpha] \text{ (undelayed)} \quad (24)$$

and

$$\cos[\omega_0(t-\tau) + \int_0^{t-\tau} \omega_I(\alpha) d\alpha] \text{ (delayed)} \quad (25)$$

where  $\omega_0$  is the desired signal frequency,  $\omega_I(t)$  is the instantaneous frequency transient, and  $t$  is the sampling time of IFM device. The measured phase shift or difference,  $\Theta$ , equals

$$\Theta = \left\{ \omega_0 t + \int_0^t \omega_I(\alpha) d\alpha \right\} - \left\{ \omega_0(t-\tau) + \int_0^{t-\tau} \omega_I(\alpha) d\alpha \right\} \quad (26)$$

which reduces to

$$\Theta = \omega_0 \tau + \int_{t-\tau}^t \omega_I(\alpha) d\alpha \quad (27)$$

which gives the frequency estimate,  $\hat{\omega}_1$ , given by

$$\hat{\omega}_1 = \omega_0 + \frac{1}{\tau} \int_{t-\tau}^t \omega_I(\alpha) d\alpha \quad t \geq \tau \quad (28)$$

The estimation bias,  $\hat{\Delta}f(t)$ , due to the instantaneous frequency transient is then written as

$$\hat{\Delta}f(t) = \frac{1}{\tau} \frac{1}{2\pi} \int_{t-\tau}^t \omega_I(\alpha) d\alpha \quad (t \geq \tau) \quad (29)$$

and

$$\hat{\Delta}f(t) = \frac{1}{\tau} \frac{1}{2\pi} \int_0^t \omega_I(\alpha) d\alpha \quad t \leq \tau \quad (30)$$

Equations (29) and (30) were programmed using the  $\omega_I(t)$  transient. The results represent the first determination of the bias error of the IFM receiver.



Because of the current interest and investment in the IFM receiver, these results may represent the most important contribution of this work.

#### ii. Applications of Analysis

Since  $\omega_T(t)$  tends to zero in the steady state, the bias error of equation (29) and (30) will tend to zero as the sampling time increase. Unfortunately, for short pulse width, the  $\omega_T(t)$  transient does not die out and the bias error may become significant.

Figure 28 plots the bias error variation in megahertz as a function of sampling time for a pulse width of 300 n.s., an offset of 20 MHz, a  $\tau$  of 10 n.s. and a filter bandwidth of 80 MHz preceding the 50% limiter. It is noted that this configuration represents the traditional IFM configuration. Referring to Figure 28, it is observed that the bias error has a damped oscillatory form with the first zero crossing occurring at twice the 25 n.s. time constant of the 80 MHz RF filter. The oscillatory nature of the bias suggests that the sampling take place at the first zero crossing. Whether this is possible operationally is a matter for further investigation. An additional run at the short pulse width of 100 n.s. is presented in Figure 29. Although deceptive in appearance, the initial bias transient is actually the same as in the 300 n.s. case which is not surprising since the initial bias is the transient response to the common leading edge.

A current trend in HPIR design is to combine the heterodyning and contiguous slot concept of the channelized receiver with the purported high accuracy of the IFM detector. In this configuration, a separate IFM receiver would be dedicated to each slot with the frequency isolation filter lowered in bandwidth

from 80 to 20 MHz. Figure 30 illustrates the effect on IFM bias of reducing the RF bandwidth for a pulse width of 300 n.s. and a signal-filter relative frequency offset of 20 MHz. It is apparent that the longer filter memory time keeps the IFM bias between 3.8 MHz and 8 MHz over the entire pulse width which is clearly an undesirable effect. It should be noted, however, that on arithmetic average over the ensemble of slots could reduce the net effect of the IFM bias to zero! The full implication of these considerations must await further investigation.

The multiple signal feature of the simulation was also used to illustrate various aspects of the strong signal capture effect. Figure 31 gives the bias error for the sum of two equal amplitude 300 n.s. signals with a relative time shift of 75 n.s. and a frequency offset of 20 MHz relative to the filter center frequency. Comparing Figure 31 to the zero time shift of Figure 30, it is observed that no significant change in  $\omega_1(t)$  occurs during the first 200 n.s. of undelayed signal; i.e., the established undelayed signal "captures" the leading edge transient of the delayed signal. Another example is given in Figure 32 where the effect of frequency offset is observed for a second 300 n.s. signal offset in frequency relative to the primary signal by 20 MHz and of magnitude one fourth that of the primary signal. The nonlinear "pulling" of the capture effect is observed in the IFM bias plot. Space constraints do not permit discussing additional interesting multiple signal effects that were observed using this flexible program.

## VI. RECOMMENDATIONS

The overall objective of the project was to optimize the High Probability Intercept Receiver. Much of the study dealt with the implications of the instantaneous frequency transient with special

attention given to the short pulse problem. In particular, the analysis and bias estimate of the IFM receiver has provided useful information on the performance of a popular receiver configuration. As expected, there are many areas that need further investigation. The following items include the more important and interesting topics for further investigation.

1. A complete investigation of the effects of random noise
2. Comprehensive comparative evaluation of the various envelope schemes and how they compare to the Crame Rao bound.
3. Investigate new methods of using the  $\omega_I(t)$  transient to reduce the short pulse problem. Among the possibilities are the use of phase lock loops and the use of the violent trailing edge transient in  $\omega_I(t)$ .
4. To investigate the effects of chirped signals and doppler on system performance.
5. To develop a "rule of thumb" for IFM operation in the presence of multiple signal effects and asymmetrical operation.
6. To investigate the analog FFT approach as a new HPIR configuration.

#### REFERENCES

1. Tsui, J.B., Microwave Receivers and Related Components, Avionics Lab, Air Force Wright Aeronautical Laboratories, PB84-108711, 1983.
2. Signorino, J.S., A Fortran Package for the Simulation and Spectral Analysis of a Channelized Receiver and Other Systems, Avionics Lab, Air Force Wright Aeronautical Laboratories, AFWAL-TR-81-1278, 1982.
3. Tsui, J.B., User Manual for REC Receiver Modeling Program, Avionics Lab, Air Force Wright Aeronautical Laboratories, Private Communication (to be published).
4. Humpherys, D.S., The Analysis and Synthesis of Electrical Filters, (Prentice Hall, Englewood Cliffs, N.J., 1970), pp. 81-87.
5. Franks, L., Signal Theory, (Prentice Hall, Englewood, Cliffs, N.J., 1969), pp. 84-85.

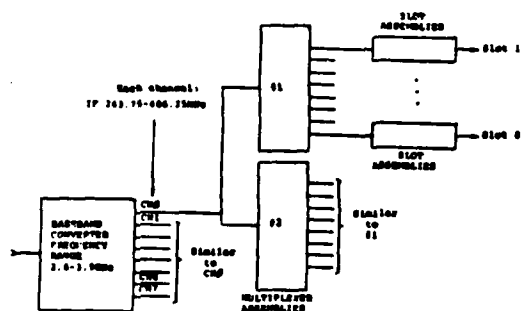


Figure 1. WPB Block Diagram of the WPB

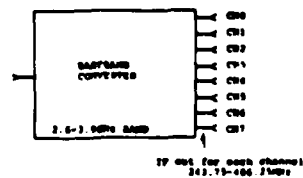
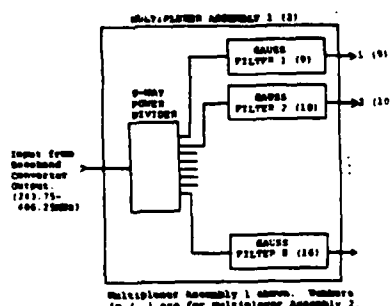


Figure 2. WPB Baseband Converter



Multiplexer Assembly 1 shown. Numbers in ( ) are for Multiplexer Assembly 2.

SLOT OUTPUT	CENTER FREQUENCY (MHz)
1 (1)	240.0 (250.0)
2 (10)	260.0 (260.0)
3 (11)	310.0 (300.0)
4 (12)	370.0 (300.0)
5 (13)	360.0 (270.0)
6 (14)	360.0 (330.0)
7 (15)	360.0 (360.0)
8 (16)	392.0 (400.0)

Figure 3. WPB Multiplexer Assembly



Figure 4: Envelope Response of First Filter (20 MHz) with Reference REC Plot.

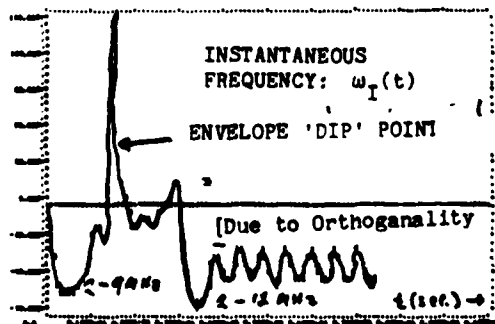


Figure 5: Instantaneous Frequency Transient Response of First Filter.

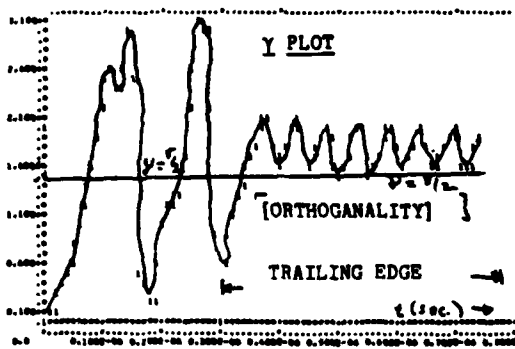


Figure 6: Orthogonality Measure,  $\gamma$ , for the First Filter

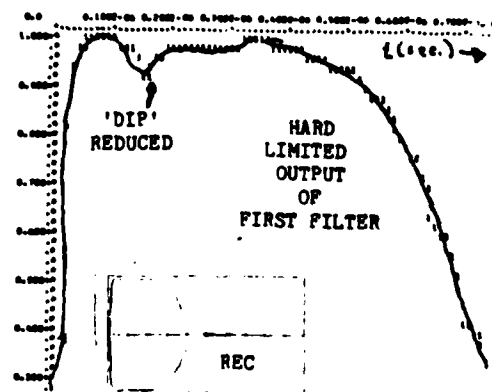


Figure 7: Envelope of Hard Limiter Output with Reference REC Plot

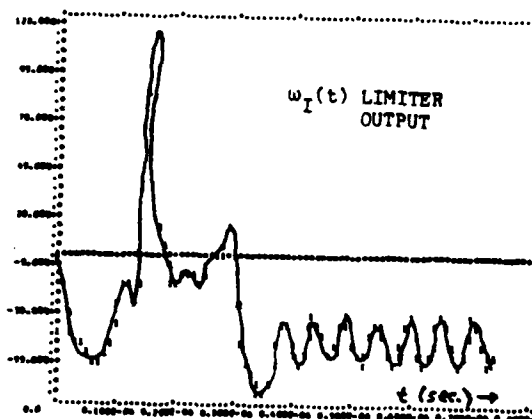


Figure 8: Instantaneous Frequency of Hard Limiter Output.

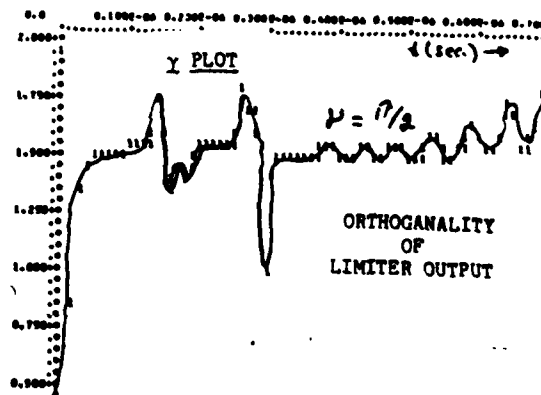


Figure 9: Orthogonality Measure,  $\gamma$ , for Hard Limiter Output.

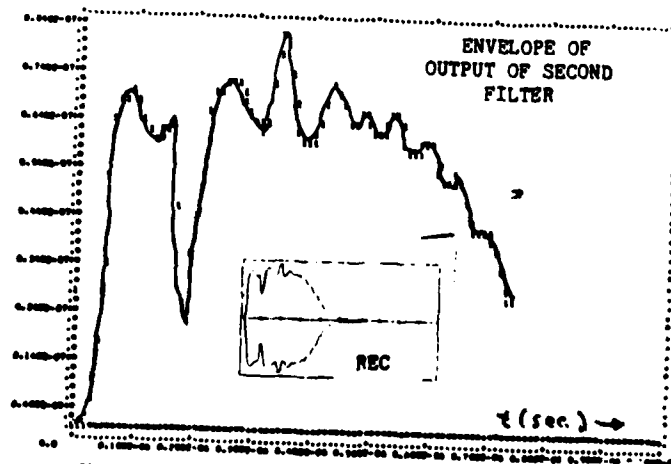


Figure 10: Envelope Response of Second Filter (40 MHz) with Reference REC Plot.

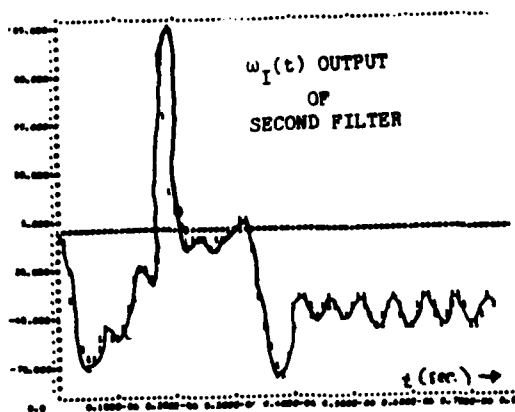


Figure 11: Instantaneous Frequency of Second Filter.

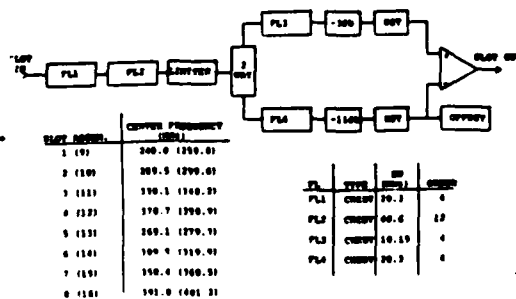


Figure 12: Schematic of Double Detection Circuit.

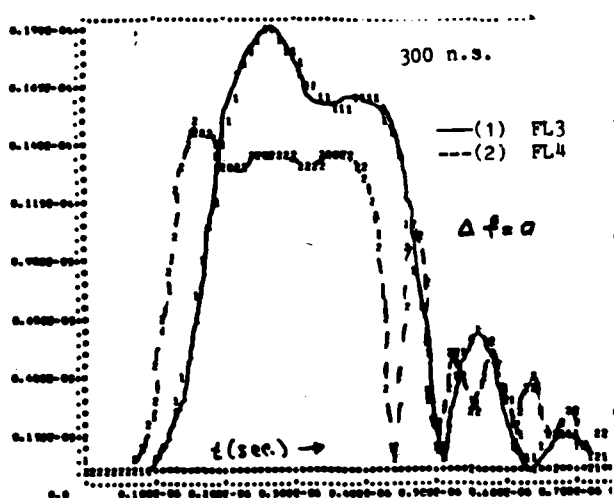


Figure 13: Double Detection Output for 300 n.s. Pulse Width With Offset of 0.

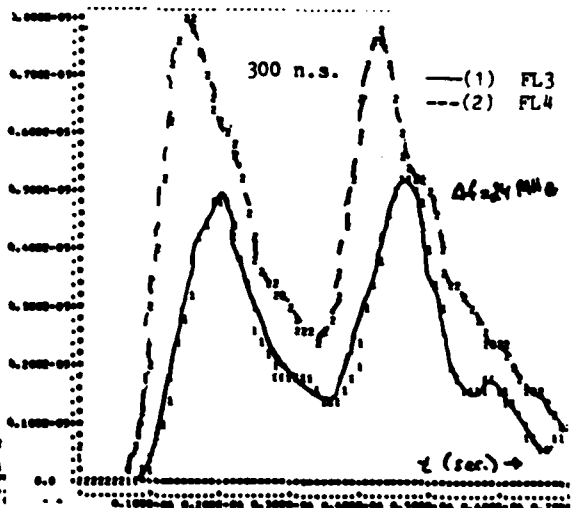


Figure 14: Double Detection Output for 300 n.s. Pulse Width With Offset of 24 MHz.

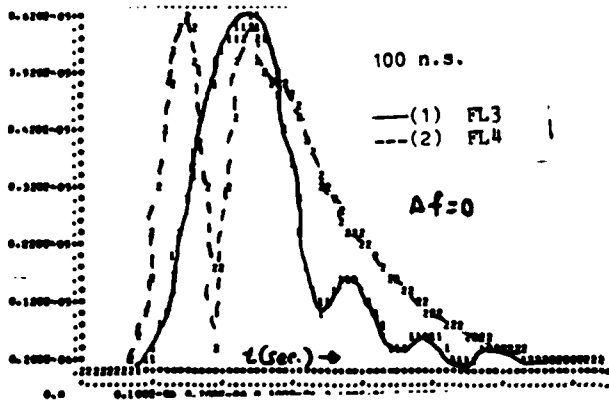


Figure 15: Double Detection Output for 100 n.s. Pulse Width With Offset of 0.

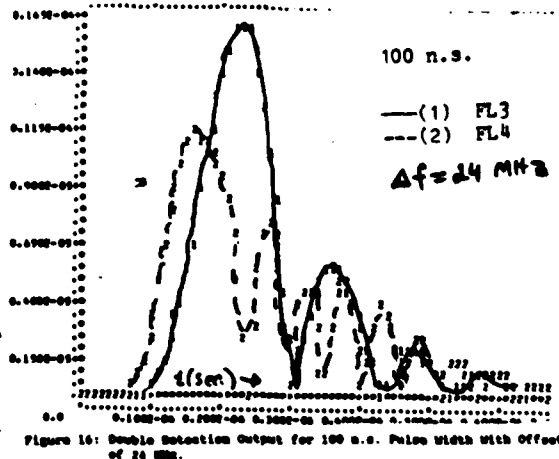


Figure 16: Double Detection Output for 100 n.s. Pulse Width With Offset of 24 MHz.

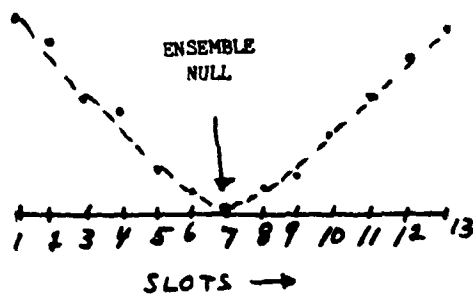


Figure 17: Ensemble Processing to Determine Minimum of Activity Index.

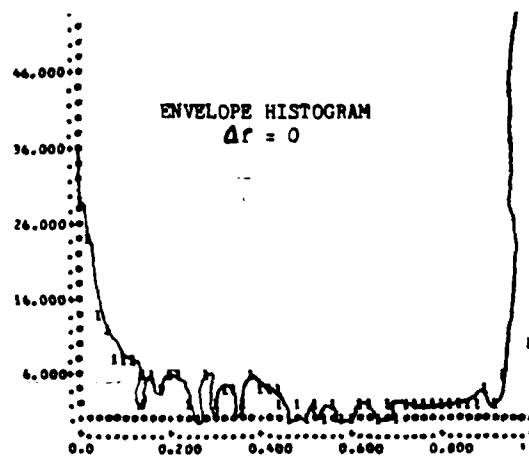


Figure 18: Envelope Histogram for 0 Offset.

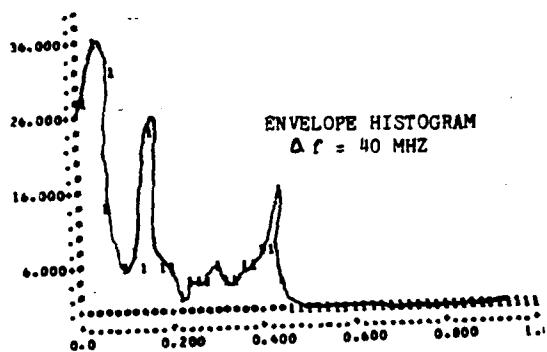


Figure 19: Envelope Histogram for 40 Mhz Offset.

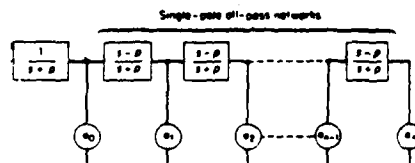


Figure 20: Laguerre Coefficient Network.

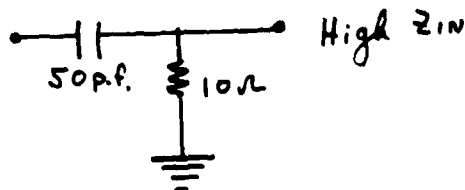


Figure 21: Load Network for FM Discriminator.

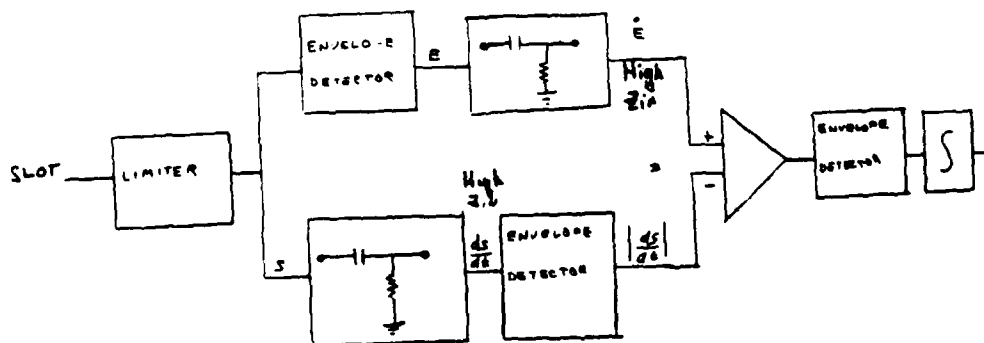


Figure 22: Schematic for FM Discriminator Scheme.



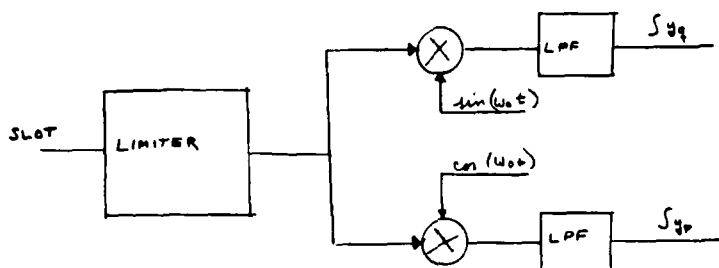


Figure 23: Quadrature Processor.

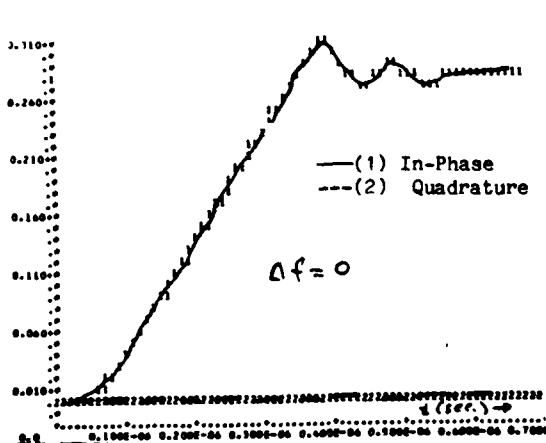


Figure 24: Integrals of In-Phase and Quadrature Components for an Offset of 0.

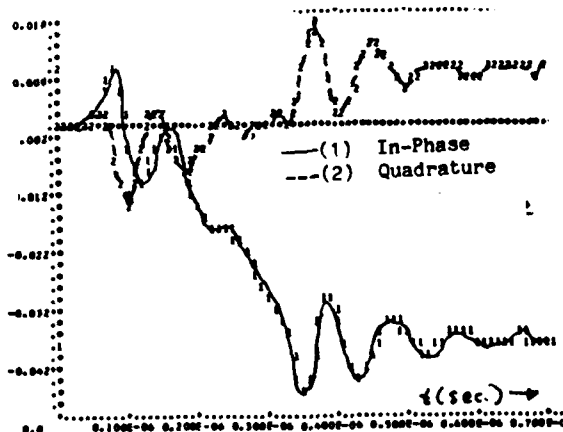


Figure 25: Integrals of In-Phase and Quadrature Components for an Offset of 24 kHz.

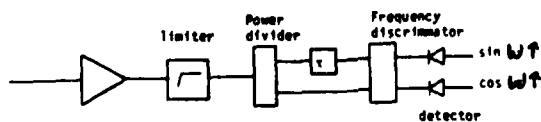


Figure 26: IFM Receiver Schematic.

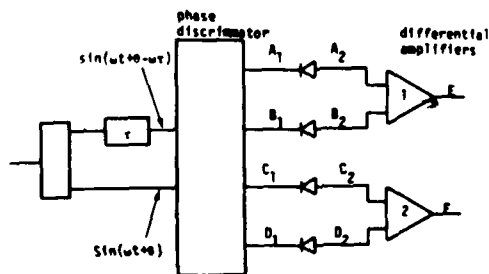


Figure 27: Detailed Schematic of IFM Phase Discriminator.

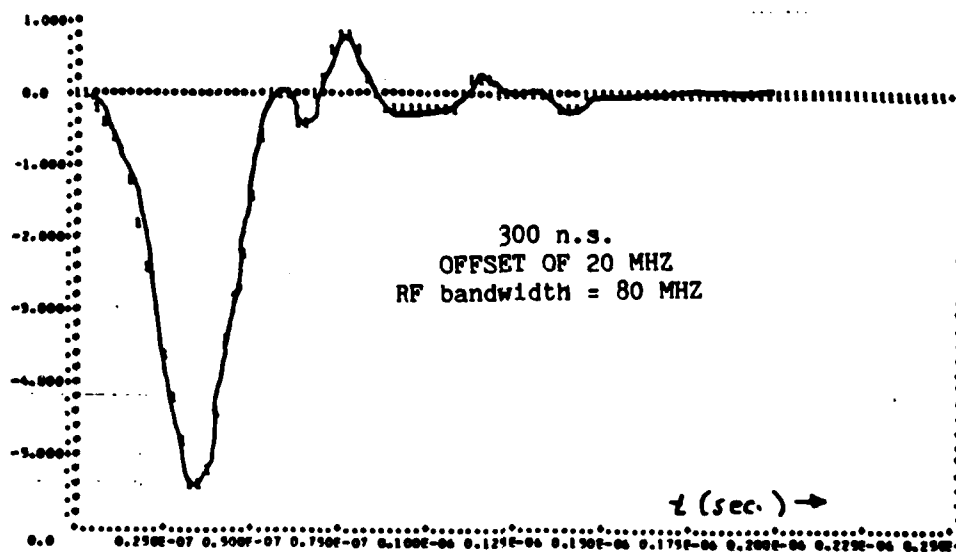


Figure 28: IFM Bias in Megahertz Versus Sampling Time for a 300 n.s. Pulse, an Offset of 20 MHz, and a RF Bandwidth of 80 MHz.

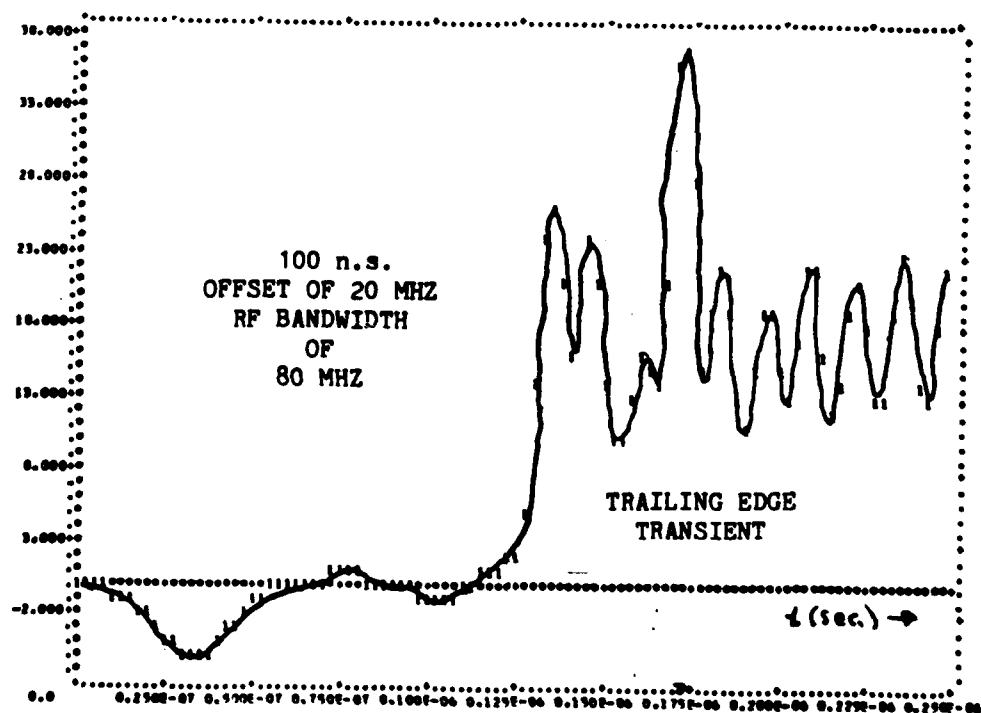
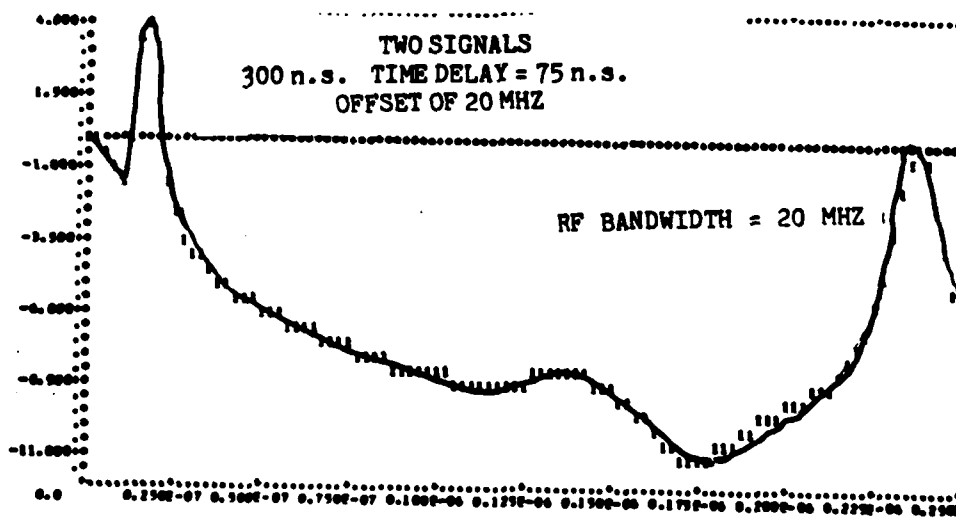
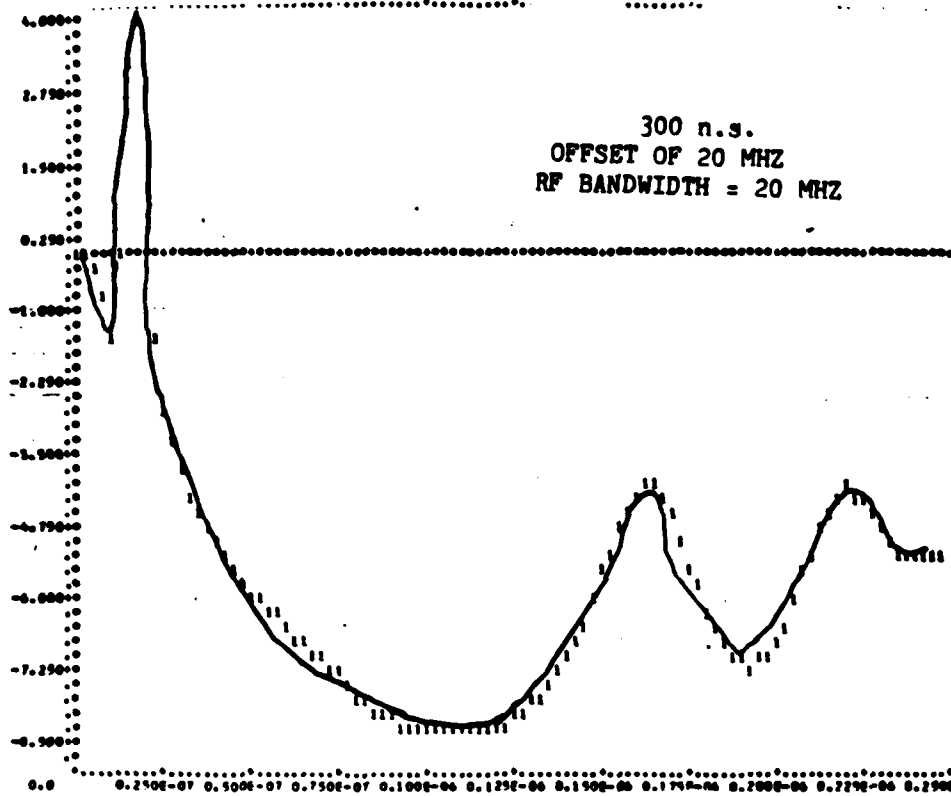


Figure 29: IFM Bias in Megahertz Versus Sampling Time for a 100 n.s. Pulse, an Offset of 20 MHz, and a RF Bandwidth of 80 MHz.



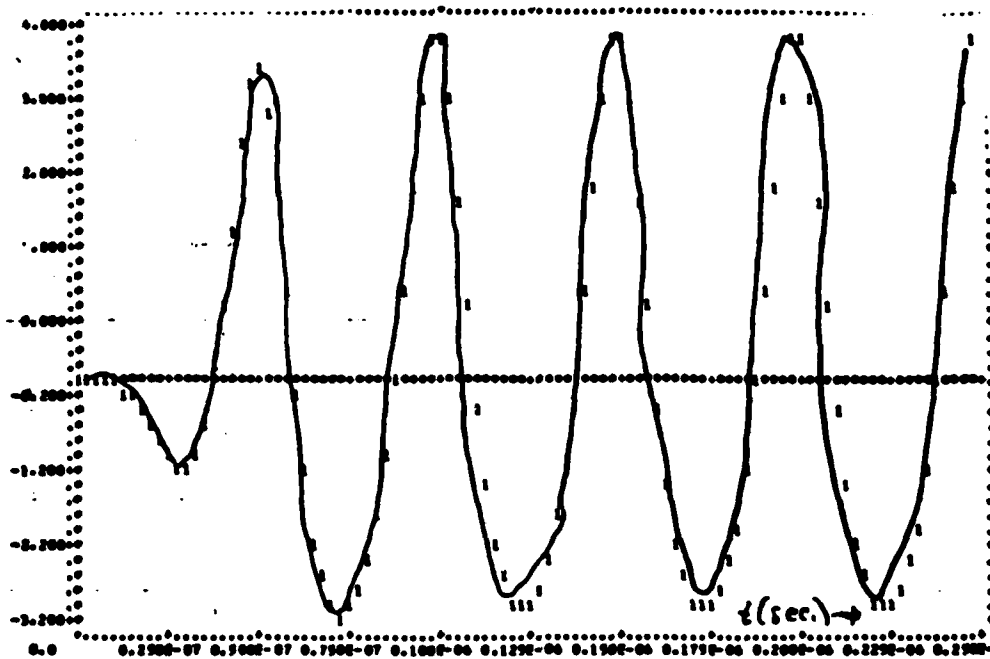


Figure 32: IFM Bias in Megahertz Versus Sampling Time for Two Signals of 300 n.s. Pulse Width, a Relative Magnitude of 4:1, and with a Mutual Frequency Offset of 20 MHz.

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

TECHNIQUES FOR ULTRA-SHORT PULSES

IN Nd:YAG LASERS

Prepared by:	Dr. Odis P. McDuff
Academic Rank:	Professor
Department and University:	Electrical Engineering Department The University of Alabama
Research Location:	Air Force School of Aerospace Medicine, Clinical Sciences Division
USAF Research	Dr. John Taboada
Date:	September 28, 1984
Contract No:	F49620-82-C-0035

TECHNIQUES FOR ULTRA-SHORT PULSES

IN Nd:YAG LASERS

by

Odis P. McDuff

ABSTRACT

The possibility of passively mode locking an actively Q-switched cw Nd:YAG laser by means of an antiresonant colliding pulse ring cavity was investigated. The gain of the medium power commercially available laser used in the study was found to be too low for mode locking to occur. Mode locking at low power was obtained when the antiresonant ring was replaced with a more conventional straight laser cavity. Harmonic doubling of the infrared energy to the visible was obtained.

An alternative scheme was suggested for the attainment of a very stable, but yet simple, mode locked laser output.

### Acknowledgement

The author would like to thank the Air Force School of Aerospace Medicine, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile and interesting summer at the Air Force School of Aerospace Medicine, Brooks AFB, Texas. He would like to thank the School, and in particular the Clinical Sciences Division, for its hospitality and provision of excellent working conditions.

Finally, the author would like to thank Dr. John Taboada for suggesting this area of research and for his many helpful comments. He would also like to acknowledge the excellent work of Mr. Heinz Jaeger of the School's Machine Shop and the excellent technical assistance of Staff Sgt. Mario Villanea.

## I. INTRODUCTION

Since the first mode-locked laser was reported in 1964,<sup>1</sup> workers in the field have produced shorter and shorter pulses and used various combinations of active and passive elements within the optical cavity for Q-switching and/or mode locking of various types of lasers.

The Air Force School of Aerospace Medicine is interested in non-invasive studies of the eye by the use of optical pulses short enough that physical mechanisms other than thermal effects come into play.<sup>2,3</sup> Stability of the system and reproducibility of the pulses is of paramount importance in work of this nature. Simultaneously Q-switched and mode-locked Nd lasers using a saturable absorber for both the Q-switcher and mode-locker produce short high-peak power pulses but tend to have large shot-to-shot variations in pulse width and energy.

The use of the colliding pulse technique<sup>4</sup> in an antiresonant ring cavity<sup>5</sup> has been shown to provide shorter and more stable pulsing in a passively Q-switched and mode-locked Nd:YAG Laser.<sup>6</sup> Purely active Q-switching and mode-locking of a cw Nd:YAG laser produced apparently stable but very low energy pulses.<sup>7</sup> A number of active-passive systems have been described<sup>8</sup> in which the Q-switching was done passively and the mode-locking done actively.

In the case of a cw Nd:YAG laser, the Q-switching must be done actively but the mode-locking could be done either passively or actively. The colliding-pulse, antiresonant cavity passive mode-locking scheme, with its greater stability over other passive mode-locking schemes and its possibility of producing shorter pulses than active mode-locking, seemed a good candidate for the production of a stable pulsed source derived from an actively Q-switched cw Nd:YAG laser.



Doubling of the output to give visible 530nm optical energy could then be accomplished for studies of the interior of the eye.

## II. OBJECTIVE

The main objective of this work was to perform an experimental investigation of the possibility of applying the colliding pulse anti-resonant ring mode-locking technique to an actively Q-switched cw Nd:YAG laser.

The laser to be used in the work was a Laser Applications Inc. Model 9560T actively Q-switched cw Nd:YAG laser which had been delivered but not put into operation. The construction of this laser, with the various components being mounted on an optical bench and removable, was such that it was a very flexible research tool.

Since the short ten-week length of the project gave no time to order and purchase items other than incidental supplies, a constraint placed upon the project was that items already available in the laboratory would have to be used. Although this affected the course of the work, it did not prevent the basic accomplishment of the project objective.

## III. THE BASIC LASER

The Laser Applications Model 9560T Nd:YAG laser had a 3mm x 79mm AR coated rod which was pumped with a Krypton arc lamp. Both were mounted in a water-cooled elliptical gold-plated pump cavity. The laser cavity mirrors were flat, one with a high reflectivity coating and the other with a 10% transmission coating on the interior surface and AR coating on the exterior surface. Removable 0.041 apertures were located at

either end of the cavity. An electromagnetic shutter and Brewster polarizer were toward one end of the cavity while a water-cooled acousto-optic Q-switcher was near the other end. The RF drive to the Q-switch was at 24 MHz and could be pulsed on and off to obtain various Q-switching repetition rates between 500 Hz and 30 KHz. A photo diode mounted on the high reflectivity mirror permitted one to monitor the temporal behavior of the Q-switched pulse. The polarizer normally was set at the factory at the polarization direction which gave maximum power output.<sup>9</sup> In this particular laser it was determined that the polarization was at about 45° from vertical. The polarizer was changed to give horizontal polarization in preparation for the antiresonant cavity work to follow. The optical length between the two cavity mirror reflecting surfaces was measured approximately as 73 cm.

The basic laser was assembled and its performance measured; generally it met the manufacturer's specifications.<sup>10</sup> Typical TEM<sub>00</sub> performance with all elements in the cavity but with the Q-switch drive off was a cw power output of 8 watts. Typical Q-switched performance gave 200 ns wide Q-switched pulses at a rep rate of 8 KHz and an average output power of 6 watts.

Looking ahead, in the use of an antiresonant ring structure to replace one of the cavity end mirrors it would be necessary to lengthen the cavity. Therefore it was of interest to determine the sensitivity of the laser performance to cavity length, since with flat mirrors it was expected that diffraction losses would depend upon the mirror locations. When the cavity was lengthened by moving the rear mirror (high reflectivity) 15 cm, holding the rear aperture fixed in position, the cw power output dropped to 2.25 watts. Also, the range of lamp

current over which lasing occurred decreased and the value of lamp current giving peak power output changed. From this experimental data,<sup>10</sup> it appeared that the laser cavity configuration was reasonably well optimized for the laser rod being used, and that modifications probably would result in poorer performance.

#### IV. THE ANTIRESONANT RING CAVITY THEORY

The antiresonant ring cavity<sup>5</sup> replaces one of the end mirrors of the laser and permits a relatively simple means of using the colliding pulse<sup>4</sup> technique to produce shorter and more stable mode-locked pulsing. The basic configuration is shown in Fig. 1.<sup>11</sup>

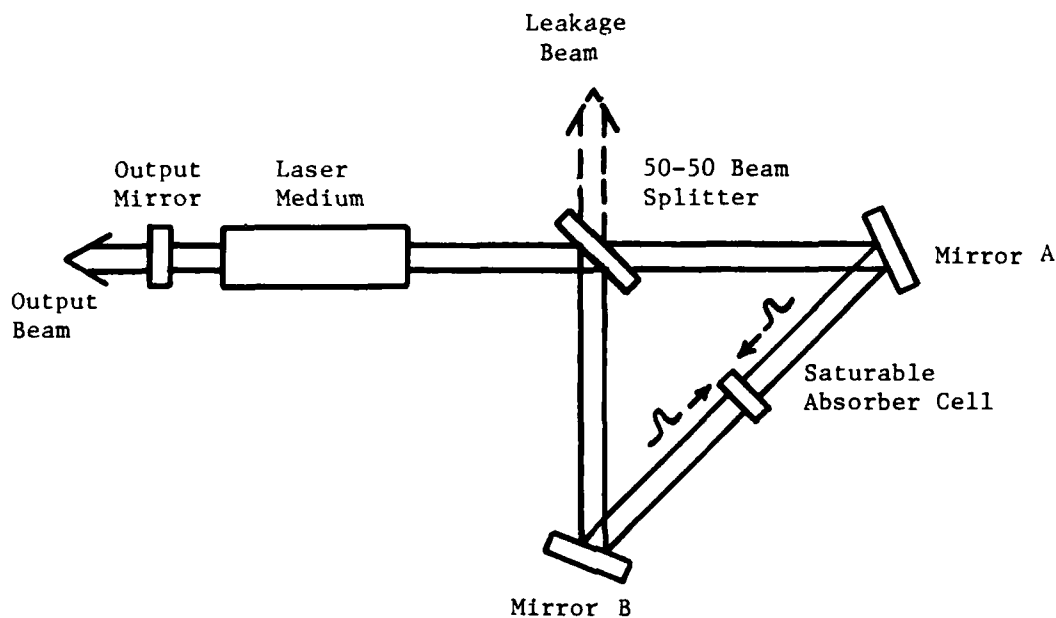


Fig. 1 Saturable-absorber mode-locking configuration using an antiresonant ring cavity.

If the beam splitter has exactly a 50-50 power division, then the signal coming into the ring from either external arm will be divided into two parts, each of which will travel around the ring in opposite directions and recombine at the beam splitter so that all the signal will be retroreflected into the same arm from whence it came. However, the clockwise (c.w.) and counterclockwise (c.c.w.) paths must be identical, otherwise there will be leakage (shown dotted in Fig. 1) and the structure will act like a partially reflecting mirror for the laser cavity.<sup>12</sup> The leakage will have a multilobed structure resulting from imperfect cancellation of the c.w. and c.c.w. beams. If the ring is symmetrical as planned, cancellation occurs for all components of the signal, pulsed or cw, assuming only that the splitting remains 50-50 at all wavelengths.

If the thin saturable absorber cell is placed exactly at the midpoint of the antiresonant ring opposite the beam splitter, as shown in Fig. 1, the two pulses will meet exactly within the absorber and the expected improved pulse formation will occur. Possible advantages of this configuration<sup>11</sup> over the conventional ring or straight double pulsing cavity<sup>4</sup> are that (1) there is only a single circulating pulse, rather than two pulses, (2) the pulse experiences full double-pass laser gain before each cell transit rather than every other transit, and (3) there is the possibility of cavity dumping using the same antiresonant cavity.<sup>13</sup>

The beam splitter does not have to be at  $45^\circ$  as sketched in Fig. 1; any physically realizable angle will do so long as the power division is 50-50. Focusing elements may be incorporated to reduce diffraction and to give better beam control. This could be done by having a curved

output mirror, or mirrors A and B could be curved. Since the reflection from mirrors A and B is at an angle, astigmatic effects will occur if they are conventional spherical mirrors. The previous work<sup>6</sup> indicates that the saturable absorber cell must be inserted at Brewster Angle to reduce the effects of multiple reflections from the cell windows.

#### V. THE CAVITY DESIGN

The beam splitter available for this work was a laser optics splitter coated for 50% reflectivity on one surface and AR coated on the other, for a  $45^\circ$  angle of incidence. This meant that the arms of the ring had to be at right angles as in Fig. 1. The mirrors available for mirrors A and B in Fig. 1 were flat and coated for high reflectivity at normal incidence. As will be seen their reflectivity was considerably lower at the  $22.5^\circ$  angle of incidence called for in the structure. The cavity (with ray paths drawn) was as sketched in Fig. 2 (not to scale).

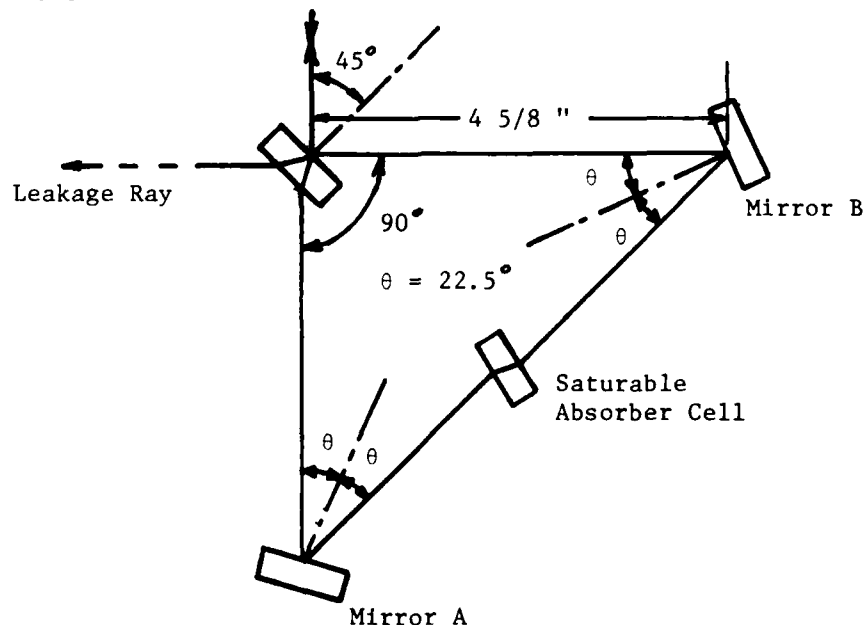


Fig. 2. The Antiresonant Cavity Design

Using a He-Ne alignment laser the beam offset due to refraction in the beam splitter was measured to be approximately  $3/32$  inch. The beam splitter and mirrors A and B all were mounted on universal mirror mounts to permit their alignment. In addition, mirror B had a translational adjustment along the axis between it and the beam splitter to compensate for the refractive offset of the beam passing through the saturable absorber cell. Thus, as the absorber cell was rotated to obtain Brewster angle, it was necessary only to translate mirror B to reestablish alignment. The absorber cell also had a vernier-adjustment of translation along the axis between mirrors A and B, since other work<sup>6</sup> had shown that the performance was critically dependent upon the cell being symmetrically located with respect to the beam splitter. The ring was made as compactly as was possible using standard available mirror mounting hardware, since making its circumference larger would lengthen the equivalent cavity and thereby increase diffraction losses.

The beam splitter and windows of the absorber cell were rotated about the beam axis until their wedging was in the plane of the ring (plane of the paper in Fig. 2). This was necessary so that the wedging effects could be compensated by adjustments of mirrors A and B about axes perpendicular to the paper in Fig. 2, thus permitting the input beam to be retroreflected back along its axis into the rest of the laser cavity.

The absorber cell was constructed of two optical flats with a teflon grommet compressed between them to give an active cell thickness of 0.8 mm. Inlet and outlet ports were drilled into one of the flats. So that the cell could easily be inserted into the ring without disturbing its previous alignment, the (slight) wedgings of the two

optical flats were oriented appositely so that they cancelled and a beam through the pair would experience minimal deflection.

The flow of the absorber dye through the cell was vertical, from bottom to top. In the closed flow supply system, the dye solution was pumped from a 165 ml reservoir supply through a filter and into the cell by means of an electromagnetic plunger (exposed parts all glass) pump developed by Dr. John Taboada and constructed by School of Aerospace Medicine shop personnel. The dye in the cell was completely exchanged in approximately two pump strokes, taking about 1 second per stroke.

A photograph of the antiresonant ring structure is shown in Fig. 3 as it appeared during the alignment process (The straight laser mirror at that end is still in place). The beam splitter, mirrors A and B, and the dye cell (white) are shown in the lower portion of the photograph, oriented as in Fig. 2. A calorimeter head is at the left of the beam splitter to measure the leakage-beam energy. A photograph from the same end of the laser showing the dye circulation system is shown in Fig. 4. The shiny case in the upper right portion of the photograph houses the power supply for the pump.

#### VI. ALIGNMENT PROCEDURE

The initial alignment of the antiresonant ring was done with the dye cell removed. As shown in Fig. 3 the entire ring structure was mounted on a baseplate which could be clamped in place at the end of the basic straight laser. Using the template of the baseplate with the light ray paths drawn in and a bore-sighted He-Ne laser beam coming from the other end as guides, the ring elements were situated to within 1/16 inch or so. Opaque discs with a mark at their center were placed in the

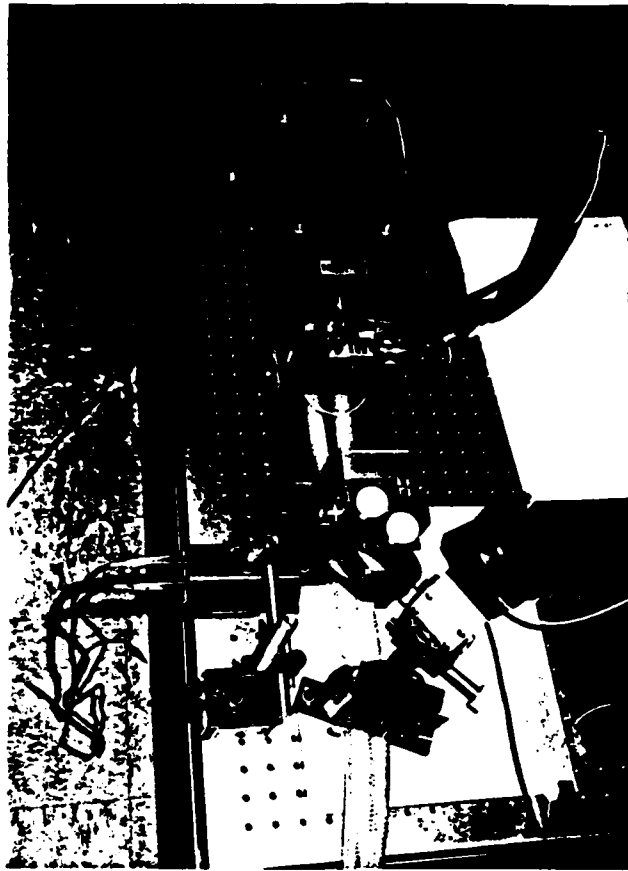


Fig. 3. Photograph of the experimental setup showing the antiresonant cavity elements.

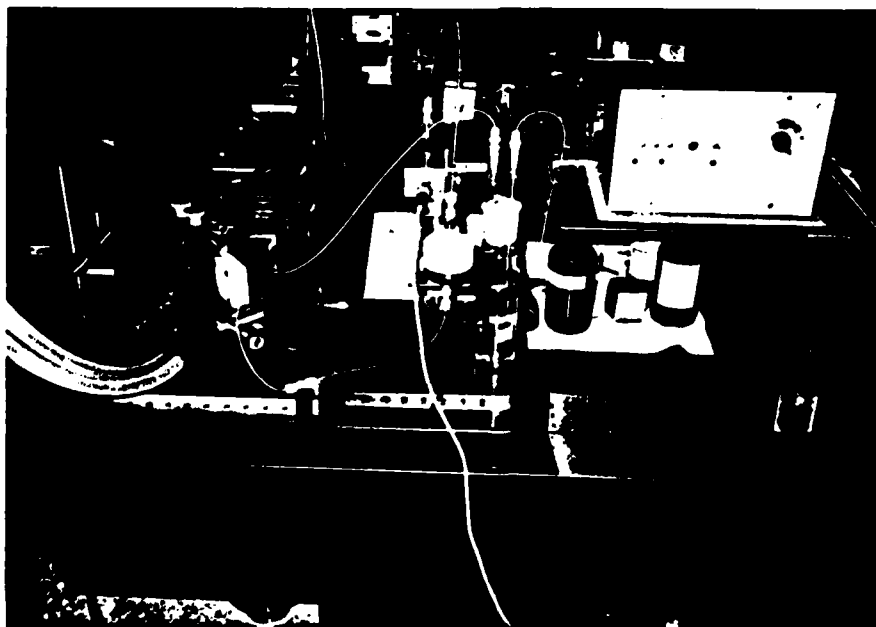


Fig. 4. Photograph of the experimental setup showing the dye circulation system adjacent to the antiresonant cavity elements.



holders for mirrors A and B and the beam splitter adjusted to place the He-Ne spot at the disc centers. Then each opaque disc in turn was removed, replacing it with its mirror, and the mirror adjusted to place its reflected spot at the center of the disc in the other corner mirror holder. Then both mirrors were placed in their holders and all elements of the ring adjusted to give retroreflection of the He-Ne beam back into the straight laser portion of the cavity. This was only an approximate alignment since in the beam splitter the refraction of the He-Ne beam is different from that of the Nd:YAG beam. In addition, it is difficult to keep track of the desirable spot since the He-Ne beam reflects almost equally well from both front and back surfaces of the corner mirrors A and B.

Then the Nd:YAG laser was operated as a straight cavity feeding directly into the antiresonant ring (with the 10% transmission mirror at the ring end and the high reflectivity mirror at the opposite end). A fluorescent infrared indicator was used to observe the leakage output and the ring elements were adjusted to give the best single mode type pattern at the leakage arm. The 10% transmission mirror was then removed and lasing occurred with the antiresonant ring acting as one of the "end mirrors". It was necessary to remove the aperture nearer to the ring in order to obtain lasing. The laser was operated in the Q-switched mode, so that output could be monitored at the high mirror reflectivity end, and the ring structure elements were adjusted to simultaneously give lowest leakage output and highest output at the other, high reflectivity end.

After the alignment was accomplished as described, and prior to insertion of the cell, an optical flat was inserted (normal to the beam)

and lasing still could be obtained without additional adjustments. Only slight tweaking of the mirrors was needed to maximize performance. As the flat was rotated toward Brewster angle it was necessary only to translate mirror B to compensate for the beam offset due to refraction in the flat. When the flat was replaced by the empty dye cell lasing could also be obtained, and the operation was essentially the same as with the optical flat.

#### VII. LASER PERFORMANCE - WITHOUT CELL

Considerable leakage was observed coming out the side arm from the beam splitter and also through the two corner mirrors, even when the structure was adjusted for optimum performance. With neither the absorber cell nor optical flat present, the leakage was as shown in Fig.5 at the lamp current of 24 amperes (slightly above the value for threshold of oscillation). The leakage through the side arm of the beam splitter was due to imperfect cancellation of the clockwise and

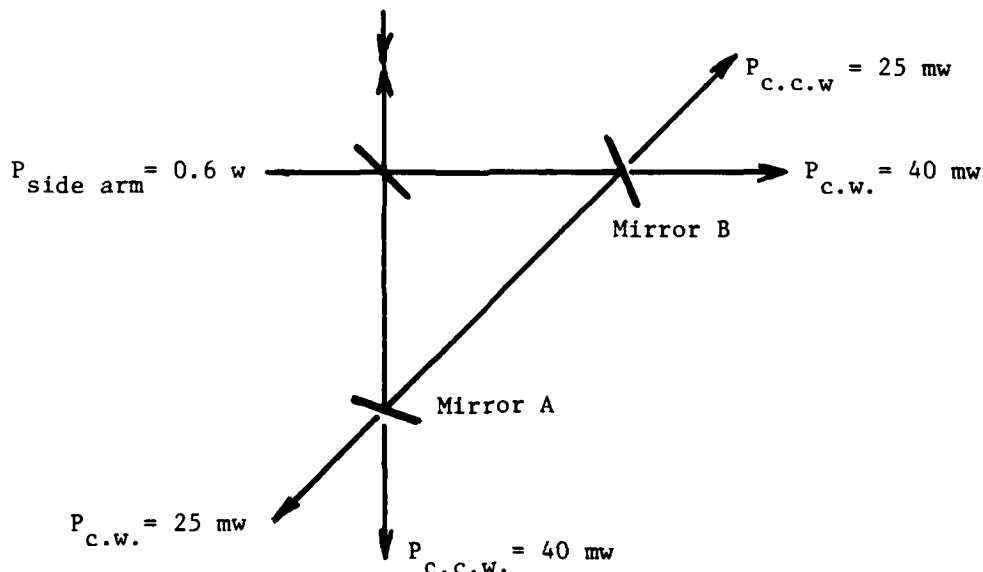


Fig. 5. Leakage powers from the antiresonant ring (high reflectivity mirror at opposite end of laser and with aperture removed from ring end). Lamp current = 24 amperes.

counterclockwise beams. It was impossible to obtain leakage much lower than this typical value. Actually, as ring adjustments were made to increase the output at the high reflectivity mirror end of the laser, the side arm beam splitter leakage increased also. The ring structure acted essentially like a relatively constant-percent transmission "mirror" at the end of the straight laser. In subsequent work the leakage arm was used as a power output port for the laser. The equality of the leakage at mirror A to that at mirror B indicated that the splitter was doing a good 50-50 split of the beam. Actually, the cancellation should be relatively insensitive to the beam splitter division ratio. If the splitter has a reflection of  $R = 0.5 \pm \delta$  and a transmission of  $T = 0.5 \mp \delta$ , the leakage power is given by the relation  $P_{\text{leakage}} = 4\delta^2 P_{\text{in}}^{12}$ . Thus a 40-60 splitter would have a leakage of only 4%. The experimental values above seemed to indicate a percent leakage much greater than that, unless the laser effective output coupling was severely undercoupled meaning that there was very high internal cavity energy storage. Replacement of the high reflectivity mirror at the opposite end by the 10% transmission mirror introduced too much output coupling and lasing was unattainable; thus, the conclusion was that the output was not undercoupled but that simply the cavity was very lossy with the ring present.

Obviously the mirrors A and B did not have high reflectivity at the  $22.5^\circ$  angle of incidence required. Later measurements external to the laser cavity indicated that their reflectivities were possibly as low as 0.6.<sup>14</sup> With the aperture at the opposite end of the laser also removed, a power output of 1.1 watts was obtained at the side arm of the antiresonant ring at a lamp current of 30 amperes.

#### VIII. LASER PERFORMANCE - WITH CELL

The insertion of the empty cell caused the power output\* from the side arm to drop by about 50%. However, the rotation of the cell to Brewster angle increased the power back up to 87% of the cell-absent value. When the cell was filled with pure chlorobenzene (to be used in the dye solution) the same performance was obtained.

The dye concentrations were prepared by adding a measured volume of a saturated solution of Kodak 9740 dye in chlorobenzene to the pumping system already filled with a known quantity of pure chlorobenzene. The initial concentration tried was one having one-fifth the dye needed to give 89% transmission through the cell. Even with this very weak concentration, the laser output power dropped to only 10 mw cw. At this low concentration mode locking could not be obtained, even with the laser Q-switched to increase the gain.

In order to obtain some confidence in the procedures being used, another high reflectivity mirror was reinserted at the ring end to give a very low-loss-cavity straight laser. The cell was inserted inside the straight cavity at the ring end. The high reflectivity mirror at the opposite end was moved to lengthen the cavity so that the  $c/2L$  mode spacing frequency would be smaller (the calculated value was 127 MHz). Then the dye concentration was increased in increments until it was five times as concentrated as the initial value above. Under Q-switched conditions mode locking was observed, but the response of the detection system was not sufficient to measure the width of the pulses.

\* Henceforth the term "power output" will refer to power out of the side arm of the beam splitter (this is actually leakage output) unless otherwise noted.

It became obvious that the gain of the laser was too low to use saturable absorber mode locking in the antiresonant cavity.

#### IX. OTHER WORK DONE

During constructional delays some effort was spent in preparation for harmonic doubling of the output of the Nd:YAG laser. A KDP doubler was constructed and the Q-switched straight laser output was doubled.<sup>10</sup> The average second harmonic power obtained was 10 mw which checked almost exactly with the theoretically calculated value of 10.8 mw.<sup>10</sup> The calculated peak power at the second harmonic was 133 watts.

Additional time was spent in checking the possibility of using active mode locking in the antiresonant cavity. An acousto-optic beam deflection crystal was available but a study of its resonances indicated that it was of the usual traveling wave type and hence not suitable for loss-modulation mode locking.<sup>14,15</sup> A check of the resonances of the Q-switcher also indicated that it was of the usual traveling wave type. A mode-locker type modulator hence was not available.

Consideration was given to the possibility of coupling the antiresonant ring to the straight mirror by means of one of the straight laser cavity mirrors. This coupled cavity approach would relax the requirements on low loss of the ring. The available 10% transmission cavity mirror did not provide strong enough coupling for this approach.

#### X. RECOMMENDATIONS

As was noted earlier the stability of the system and reproducibility of the pulses is of paramount importance in using a short pulsed laser for non-invasive studies of the eye. Although the

combination of active Q-switching and passive mode locking in an antiresonant ring cavity promised to be an improvement over previous passive-passive or passive-active systems, a more efficient cavity structure or a higher gain laser than available in this work will be required. The active-passive, antiresonant scheme based on the previous work with non-cw Nd:YAG lasers<sup>6</sup> still may not be stable enough. Commercial active-active systems are becoming available which use exotic temperature compensation techniques in stabilizing the laser cavity and the mode-locker individually<sup>16</sup> but they are quite expensive.

It is recommended that consideration be given to an active-active system, followed by harmonic doubling and pulse compression techniques. It should be possible to develop an active-active system where the mode locker frequency is coupled to the cavity  $c/2L$  frequency, letting both drift about, as long as they stay together. Past techniques at active stabilization have attempted to hold the laser  $c/2L$  constant<sup>17,18</sup> independently and then hold the modulator frequency at this value.

The requirement for an active-active system where the modulation frequency is locked to the  $c/2L$  frequency is that it must be able to sense when the modulator frequency departs from the  $c/2L$  mode spacing. Such a discriminant has been shown theoretically to exist for loss-modulation active mode locking,<sup>19</sup> and experimentally for phase-modulation active mode locking.<sup>20</sup> Basically, the phase of the mode locked pulse shifts with respect to the modulator drive signal when the drive frequency departs from the  $c/2L$  frequency. It is recommended that this concept be given further study. It should be possible to develop a very stable, relatively simple mode-locked laser which avoids all the complicated temperature compensation schemes.

### REFERENCES

1. L. E. Hargrove, R. L. Fork and M. A. Pollack, "Locking of He-Ne Laser Modes Induced by Synchronous Intracavity Modulation," Appl. Phys. Letters, Vol. 5, pp. 4-5, July 1964.
2. J. Taboada and W. D. Gibbons, "Retinal Tissue Damage Induced by Single Ultrashort 1060 nm Laser Light Pulses," Applied Optics, Vol. 17, pp. 2871-2873, Sept. 1978.
3. W. T. Ham, Jr., H. A. Mueller, A. I. Goldman, B. E. Newman, L. M. Holland and T. Kuwabara, "Ocular Hazard from Picosecond Pulses of Nd:YAG Laser Radiation," Science, Vol. 185, pp. 362-363, July 1984.
4. R. L. Fork, B. I. Greene and C. V. Shank, "Generation of Optical Pulses Shorter than 0.1 psec by Colliding Pulse Mode Locking," Appl. Phys. Letters, Vol. 38, pp. 671-672, May 1981.
5. A. E. Siegman, "An Antiresonant Ring Interferometer for Coupled Laser Cavities, Laser Output Coupling, Mode Locking, and Cavity Dumping," IEEE J. Quantum Electronics, Vol. QE-9, pp. 247-250, Feb. 1973.
6. H. Vanherzeele, J. L. Van Eck and A. E. Siegman, "Colliding Pulse Mode Locking of a Nd:YAG Laser with an Antiresonant Ring Structure," Applied Optics, Vol. 20, pp. 3484-3486, Oct. 1981.
7. D. J. Kuizenga, D. W. Phillion, T. Lund and A. E. Seigman, "Simultaneous Q-Switching and Mode-Locking in the CW Nd:YAG Laser," Opt. Commun., Vol. 9, pp. 221-226, Nov. 1973.
8. M. A. Lewis and J. T. Knudtson, "Active-Passive Mode-Locked Nd:YAG Oscillator," Applied Optics, Vol. 21, pp. 2897-2900, Aug. 1982. (This paper has a number of references to active-passive systems.)
9. Personal communication from Mr. Doug King of Laser Applications, Inc.
10. See Laboratory Notebook, Vol. 1, of the author.
11. A. E. Siegman, "Passive Mode Locking Using an Antiresonant-Ring Laser Cavity," Optics Letters, Vol. 6, pp. 334-335, July 1981.
12. S. C. Sheng and A. E. Siegman, "Perturbed Gaussian Modes of an Unbalanced Antiresonant-Ring Laser Cavity," J. Opt. Soc. Am., Vol. 66, pp. 1032-1036, Oct. 1976.
13. R. Trutna and A. E. Siegman, "Laser Cavity Dumping Using an Antiresonant Ring," IEEE J. Quantum Electronics, Vol. QE-13, pp. 955-962, Dec. 1977.

14. See Laboratory Notebook, Vol. 2, of the author.
15. R. T. Denton, "Modulation Techniques," Chapter C6 of Laser Handbook Vol. 1, F. T. Arecchi and E. O. Schulz-Dubois, eds., North Holland Publishing Co., Amsterdam, 1972.
16. See, for example, advertising brochure for Spectra Physics series 3000 Nd:YAG lasers.
17. O. Svelto, "Techniques of Solid-State Lasers," Chapter C2 of Laser Handbook, Vol. 2, F. T. Arecchi and E. O. Schulz-Dubois, eds., North Holland Publishing Co., Amsterdam, 1972.
18. O. P. McDuff, "Techniques of Gas Lasers," Chapter C5 of Laser Handbook, Vol. 2, F. T. Arecchi and E. O. Schulz-Dubois, eds., North Holland Publishing Co., Amsterdam, 1972.
19. O. P. McDuff and S. E. Harris, "Nonlinear Theory of the Internally Loss Modulated Laser," IEEE J. Quant. Electronics, Vol. QE-3, pp. 101-111, March 1967.
20. L. M. Osterink and J. D. Foster, "A Mode-Locked Nd:YAG laser," J. Appl. Phys., Vol. 39, pp. 4163-4165, August 1968.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

PLASMA GENERATION AND DIAGNOSTICS FOR IONOSPHERIC PLASMA SIMULATION

Prepared by:	Dr. Bernard McIntyre
Academic Rank:	Assoc. Professor
Department and university	Department of Computer Engineering Technology University of Houston, University Park
Research Location:	Air Force Geophysics Laboratory, Hanscom AFB, Mass.
USAF Research	Herbert A. Cohen
Date:	August 20 1984
Contract No:	F49620-82-C-0035

PLASMA GENERATION AND DIAGNOSTICS FOR IONOSPHERIC  
PLASMA SIMULATION

by

Bernard McIntyre

ABSTRACT

An electron/ion source was tested for possible use as a plasma source in ionospheric simulation where a plasma density of  $10^3 - 10^5$  /cc is required. A system of Langmuir probes was used to monitor electron and ion concentrations in the Jumbo vacuum chamber of the Air Force Geophysics Laboratory and showed that far from the source the plasma density is only  $10^3$  /cc and that there are very likely significant electric fields in the chamber due to the large Debye length associated with this plasma. Small ratios of electron to ion current densities were observed and attributed to the presence of the electric fields.

#### Acknowledgement

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to interact with an active research program at AFGL. He would especially like to thank Herbert A. Cohen and the members of his group for all of the help and cooperation extended to him during the summer tenure.

## I INTRODUCTION

The AFGL is preparing for a test in which electron and ion sources will be operated on a vehicle in order to study spacecraft charging, beam propagation, and electromagnetic interference. It is expected that the vehicle will encounter in its flight an ionospheric plasma having a number density ranging from  $10^3$  to  $10^5$  /cc. A complete ground test of the system will be conducted in a large vacuum chamber in which the ionospheric conditions should be simulated as closely as possible. In addition, individual components of the payload will be tested in smaller vacuum chambers at the AFGL. In order to perform the plasma simulation at these two extremes a plasma source capable of simulating the required charge background over a wide range of test volumes is required. In addition, a system of plasma diagnostics is required to be able to monitor the plasma parameters quickly and accurately. This paper summarizes the work done by the author at the AFGL on plasma sources and diagnostics and presented in a series of talks to the Space Physics Division of the AFGL during the summer. The subject of ion collection is studied here using the sharp sheath model of Langmuir and Mott-Smith<sup>(1)</sup>. Deviations from their predicted ratio of electron to ion current are noted and attributed to the probe being in the sheath formed at the wall of the plasma boundary.

## II OBJECTIVES

The main objectives of the research effort were to develop techniques for simulating ionospheric conditions in vacuum chambers and design a portable plasma diagnostics system which

can be used to monitor the plasma parameters. Data would be required on the electron and ion densities, temperatures and plasma potentials.

### III LANGMUIR PROBE THEORY

The theory of biased electrostatic probes in a collisionless plasma was developed by Langmuir and Mott-Smith<sup>(1)</sup> and correctly predicted the relationship between the plasma current to the probe versus the bias potential for a wide range of plasmas. A typical current-voltage characteristic is shown in Figure 1., in which the voltage axis represents the voltage referenced to the chamber ground.

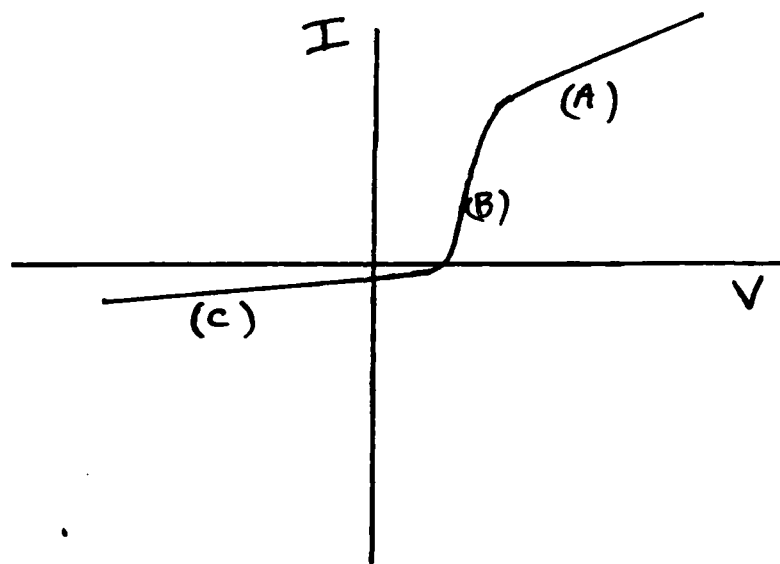


Figure 1. Current-voltage characteristic for a biased probe in a plasma.

In region (A) of the characteristic the biased probe is positive relative to ground and collects electrons. For a low density collisionless plasma, in which the collected charges undergo no collisions while in the electric field of the probe, Langmuir

and Smith showed that the electron part of the curve varies as  $V^1$  for a sphere and  $V^{1/2}$  for a cylinder. In region B, the electrons are being repelled from the probe and the curve is exponential for Maxwellian electrons. The current in an exponential region varies as

$$I = I_p \exp \left[ e(V - V_p) / K T_e \right] , \quad (1)$$

where  $V_p$  is the plasma potential,  $K$  is the Boltzmann constant and  $T_e$  is the electron temperature. The point  $V$  is found by using a semilog plot of the current-voltage profile and serves as the plasma ground in that the probe is not disturbing the plasma when at that potential. When  $V > V_p$  electrons are attracted to the probe and when  $V < V_p$  electrons are repelled and ions attracted. When the probe is at  $V_p$  the ambient plasma electrons and ions which intersect the probe are collected, assuming that there is no reflection or secondary electron emission at the probe surface. The ratio of electron to ion current at  $V_p$  then should scale as  $(M_i / M_e)^{1/2}$  where  $M_i$  and  $M_e$  are the ion and electron masses. Temperature effects are not usually considered important in this current ratio for reasons which will be considered below when discussing sheath formation.

For section (C) of the probe characteristic, the probe is negative relative to the plasma and all electrons are repelled by it. The probe theory of Langmuir and Mott-Smith makes no distinction between ion collection and electron collection in that one simply substitutes the proper mass to calculate

currents. The ion collecting part of the curve then should also vary as  $V^{1/2}$  for a cylindrical probe in a collisionless plasma. Bohm<sup>(2)</sup> later showed that the slower of the charged particles, usually the ions, must enter the collecting region of the probe with about one half of the thermal energy that the electrons have. Using this assumption, the ion part of the curve can be extrapolated back to  $V_p$  and equated at that point to the random thermal current of ions having an energy of about  $(1/2)KT_e$ . The ion current derived by Bohm is given by

$$I_+ = 0.4N_i e (2KT_e / M_i)^{1/2} , \quad (2)$$

where  $N_i$  is the ion density. The electron temperature is found from the exponential section, (B), by calculating the slope,  $1/KT_e$ , in a semilog plot. Once  $KT_e$  is known, equation (2) can be used to calculate  $N_i$ . A similar approach can be used to calculate  $N_e$  since the electron current at  $V_p$  is given by

$$I_- = N_e e V_{thermal} / 4 \quad (3)$$

When using equations (2) or (3), care must be taken to subtract the effects of ion currents from electron current data points and vice versa. For dense plasmas ( $N > 10^6$ /cc) this becomes a problem only in the region where  $I \approx 0$  and the probe collects both electrons and ions. Near  $V_p$  for instance, the ion contributions in the case of argon should be about  $10^{-2}$  times that of the electrons if the above theory is applicable. In

practice, the main difficulty with using equations (2) and (3) to calculate  $N_i$  and  $N_e$  independently is that the surface effects on the ion collecting probe are usually uncertain. Metastable gas atoms, for example, could bombard the surface and release electrons. Ion bombardment of the probe causes secondary electron emission as the voltage increases and the ideal ion probe characteristic may not be obtained. This makes the extrapolation process more dubious also. Usually then, the electron part of the characteristic is used to calculate  $N_e$  and it is assumed  $N_i \approx N_e$ . The ion part can then be used as a guide or rough check on the value of  $N_e$ .

#### IV SHEATH THEORY

A better understanding of the applicability and limitations of the Langmuir probe theory of section III can be found from the study of sheath formation. According to Langmuir, a biased probe will repel one charge species and collect the other. If the probe potential,  $V$ , is much greater than  $KT/e$ , all the charges of one type are turned back in a distance over which  $\Delta V \approx KT/e$ . Assuming  $T_i \ll T_e$ , the ions are turned back very easily and for simplicity, Langmuir assumed a sharp sheath edge at which  $V = V_p$ . At this edge only electrons penetrate, entering at their thermal energy,  $KT_e$ . Langmuir used the same sharp sheath edge concept for ion collection. When  $V \gg KT_e/e$  and the region of space over  $\Delta V \approx KT/e$  (and the electrons are turned back) is negligible, ion collection and electron collection can be treated on an equal basis in that current collection is modified by the mass and temperature of the charges. The biased probe is surrounded by a space charge, positive or negative, which shields the probe from



the ambient plasma in which  $V = V_p$  and the region between the probe and the plasma is the sheath region.

(2) Bohm later showed that sheath stability can exist only when the slower charge is accelerated from beyond the sheath edge in a zone called the pre-sheath. The charge is accelerated in the pre-sheath so that it enters the sheath with an energy  $\geq (1/2)KT$  where  $KT$  is the thermal energy of the faster charge. Ion collection can be complicated then by two important factors. First, the sheath edge at which the electrons are turned back can be very fuzzy or thick. Second, there is a hard to define region called the pre-sheath in which  $N_i \approx N_e$  and ion motions are directed to the probe and the ions gain energy. If the plasma can not be considered collisionless or if there are other electric fields present in the pre-sheath region, an ion sheath may not form and the ion part of probe curve may not be of use in finding  $N_i$ .

#### V DEBYE LENGTH

The Debye length is a useful concept for discussing some of the above problems in ion collection. (3)

If a point charge is maintained at a point in an ionized gas in which there are roughly an equal number of electrons and ions, the point charge gets shielded out by the opposite charge so that the potential relative to the fixed point charge decreases faster than  $1/r$ . In the steady state and for potentials small compared to  $KT_e$ , the potential varies as

$$V(r) = \frac{V_0}{r} \exp(-r/\lambda_D) \quad , \quad (4)$$

where  $\lambda_D = 69(T_e / N_e)^{1/2}$  ( $T = ^\circ K$ ;  $N_e = \# / m^3$ ) is the Debye length or shielding distance. Eq (4) can be thought of as the potential due to a small charge or the form of the potential on the fringe of a field due to a large charge accumulation.

In the case of a probe collecting ions, the pre-sheath potential is  $\Delta V \approx (KT_e)/2$  and the pre-sheath region would be expected to be at least a Debye length in extent. For  $KT_e = 1$  ev and  $N_e = 10^4/cc$ ,  $\lambda_D = 7.4$  cm. When  $\lambda_D$  is not small compared to the size of the plasma container or particle mean free paths, the sheath theory of Langmuir is not applicable. In the latter situation, it is questionable whether an ion sheath ever forms about the collector and more sophisticated calculations must be performed. One more point should be mentioned in discussing the applicability of the sharp sheath edge theory to probe collection. The Langmuir - Smith relation  $I \propto V^{1/2}$  results from considering angular momentum effects to charges which enter a sharp sheath edge. Current collection is orbit limited and  $\propto V^{1/2}$  when the sheaths are so large that some charges enter the sheath but are not collected. For all of the data collected for this paper the current collection is orbit limited and the key issue related to the collection of ions is whether or not sheath formation can be assumed.

## VI PROBE DATA ANALYSIS

A system of Langmuir probes was used to study the plasma generated by a plasma source from the Spire Corp. This source consisted of a wire mesh cathode inside a cylindrical anode which could be floated relative to an extractor grid in front of the

anode as shown in figure 2.

At a gas flow of 1 - 4 SCC/min the electrons from the cathode bombard the gas atoms and create electron-ion pairs near the anode aperture. Since electrons are attracted to the anode they can be extracted via the electron sheath that forms near the anode. Initially, probe data was taken with the anode grounded and the extractor power supply off and the extractor grid grounded. The cathode was then about -40V relative to the anode. This configuration was used so that a probe could be inserted between the cathode and the anode without having the probe system, shown in Figure 3, floating at a high voltage. The internal probe data is shown in Figure 4, from which  $N_i = 1.3 \times 10^7 / \text{cc}$ ,  $V_p = -25\text{V}$  and  $KT_e = 1.8\text{eV}$ . Note that  $V_p = -25\text{V}$  relative to chamber ground since the anode is grounded. The pressure in the source is estimated at approximately .5 Torr based on the gas flow and source aperture size. The Debye length is 0.3 cm, and the mean free paths  $\lambda_e = 0.03$  cm and  $\lambda_i = 0.01$  cm. As expected, the cathode-anode region is collision dominated and for this relatively high density case the ratio of electron to ion current near  $V_p$  is

$$I_- / I_+ = \frac{1.8 \times 10^{-6} \text{ amp}}{4 \times 10^{-9} \text{ amp}} = 450, \quad (5)$$

and is about twice the theoretical value. The important points to note for this data are that the ratio of  $I_- / I_+$  is very reasonable even though ion and electron collection are apparently effected by collisions as evidenced by the fact that neither  $I_+$

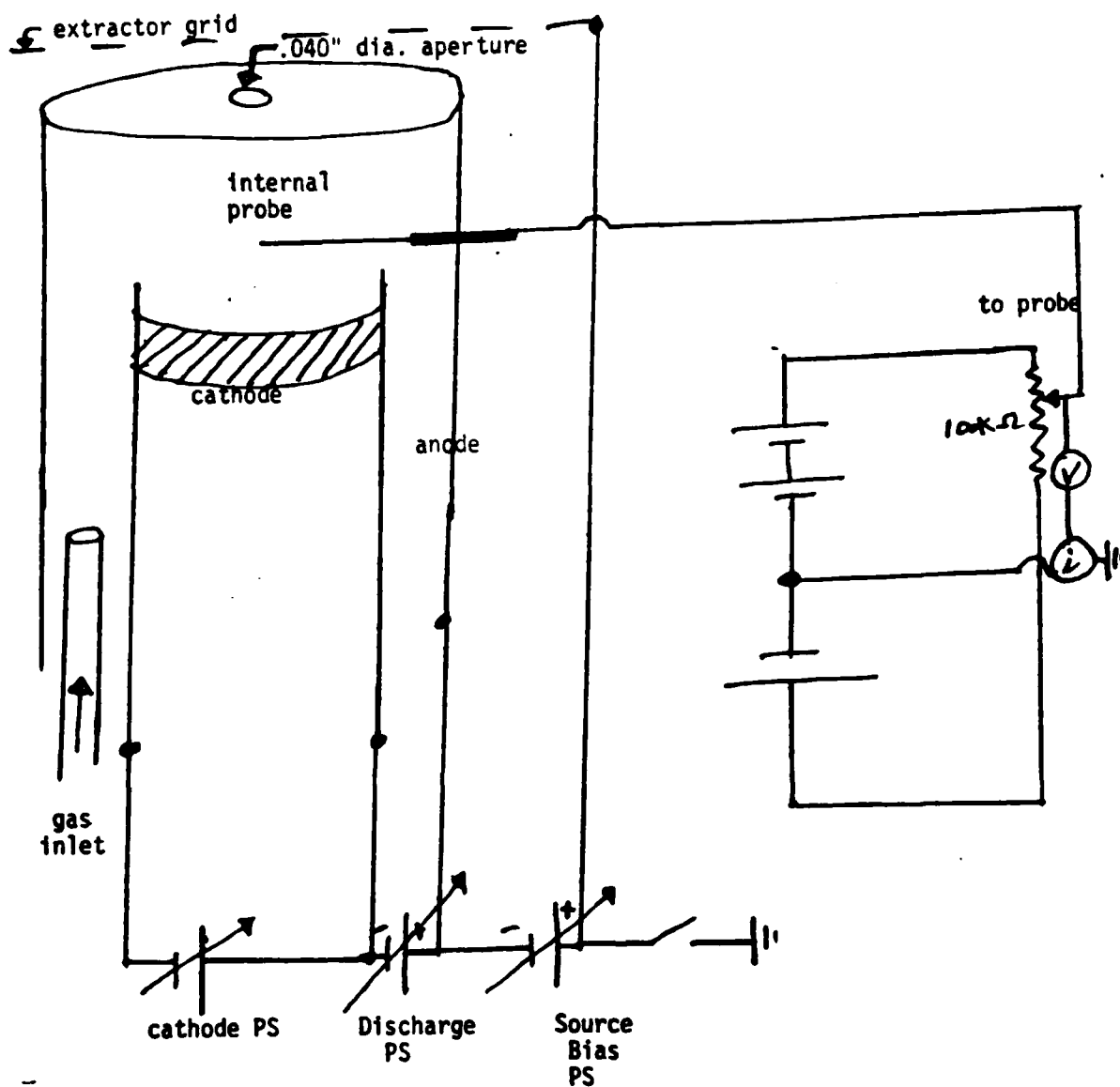


Figure 2. Spire plasma source and power supply configuration

Figure 3. Langmuir Probe System

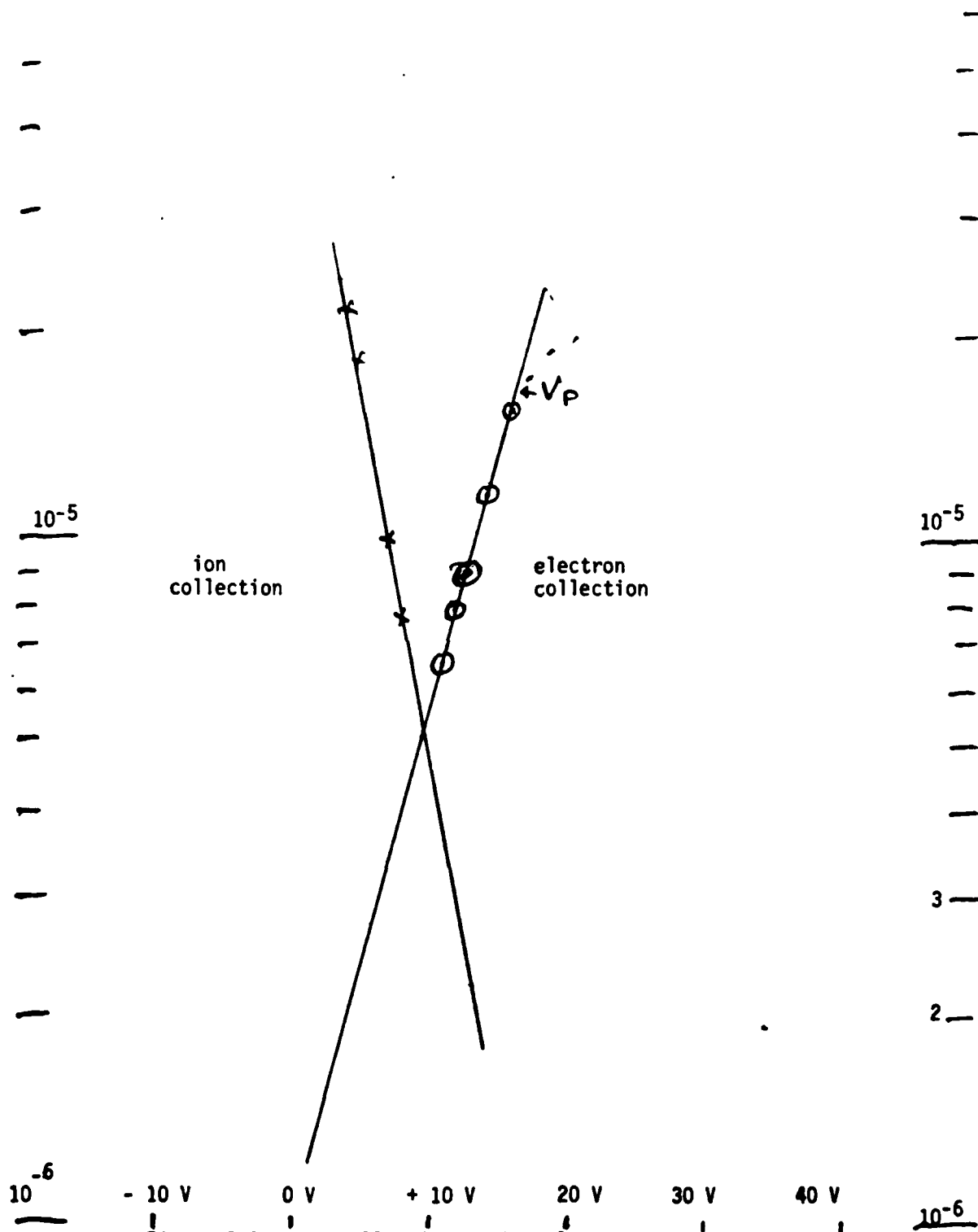


Figure 5 Current collection inside plasma source

nor  $I_-$  vary as  $V^{1/2}$ . Also, since  $V_p = -25V$  and the anode is at ground while the cathode is at  $-40V$ , the probe is not in the sheath of the cathode or the anode.

The plasma source was then used with the extractor power supply on but with all parts of the source floating. Ground connection could only occur then via the plasma external to the source. The bias power supply between the anode and extractor grid was set at  $320V$  and the grid was observed to float to  $+400V$  relative to ground. In this way the grid could collect electrons from the plasma for its current return loop. The internal probe characteristic, shown in figure 5, indicated  $N_e = 5.6 \times 10^7 / cc$ ,  $KTe = 5.5eV$  and  $V_p = +17$  while the anode was  $+10V$ , all relative to chamber ground. The larger value of  $N_e$  was due to a higher gas flow. The ion curve was completely different from the previous case though as the ion current was approximately equal to the electron current near  $V_p$ . This implies that either  $N_i \gg N_e$  or the ions have extremely high thermal energy. The author concludes here that  $N_i \gg N_e$  and the probe is likely in the wall sheath of the anode since  $V_p$  is only two volts from the anode potential. The effect of the wall sheath is to contain the plasma and keep the faster charge from depleting the volume; in effect, it tends to force  $i_+ = i_-$ .

This floating source configuration, in which the intent was to simulate charge emission from an isolated satellite, indicated that the outer grid, analogous to the spacecraft ground, extracted electrons out of the source interior but then floated at about  $+100V$  relative to the chamber plasma so that it could attract plasma electrons back to it. The anode also floated positive,

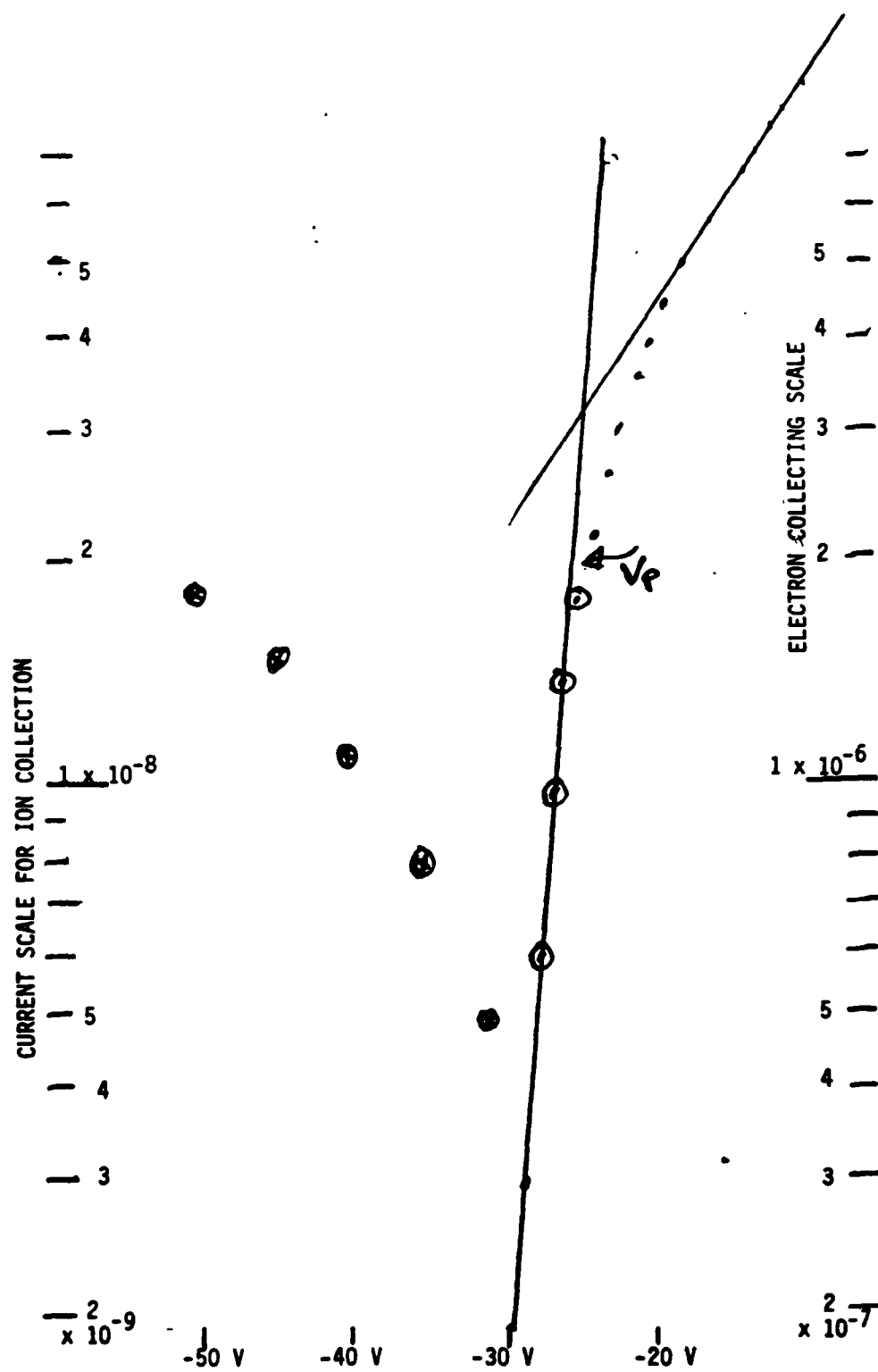


Figure 4. Current collection for probe inside plasma source

about +20V, since it was exposed to the chamber plasma. It would have been interesting to have had a second independent plasma source to test the behavior of the floating grid under various plasma conditions.

Several probe characteristics were taken in the center of the vacuum chamber under various operating conditions of the source. In all of these situations  $N_e$  was in the range  $1.5 \times 10^3$  /cc,  $\lambda_D = 10$ cm and  $I_+ \approx I_-$ . Since the chamber wall sheath extends for at least one Debye length and the chamber diameter  $\approx 1$ m, the volume enclosed by the wall sheath is not small compared to the remainder of the chamber volume, and the dominant condition is  $I_+ \approx I_-$  rather than  $N_i \approx N_e$ . Ratios of  $I_+/I_-$  have been noted by Chen and Goodings.<sup>(4)</sup> They observed a trend toward higher  $i_+/i_-$  ratios as the plasma density decreased below  $10^7$ /cc. They concluded that the ion sheath probably is not well defined for these large Debye length and  $N_i \neq N_e$ .

## VII RECOMMENDATIONS

The data in section VI indicates that there are three areas in which the above work could be extended.

First, when the source was used in a floating configuration, isolated from chamber ground, the outer grid or extractor was set at +320V relative to the anode. Electrons were then extracted from the anode and exited through the extractor grid. As a result, the source went to a positive voltage and then attracted electrons back from the plasma in the chamber. The chamber plasma had resulted from the source electrons ionizing gas atoms outside the source. The extractor grid in this case is used as



an acceleration/deceleration grid. The grid required a bias of +80V relative to the plasma in order to collect enough electrons to balance its electron losses as a source. A second plasma source could be used to induce the charging and discharging of the grid as the background plasma density is lowered and raised. Using a small electron or ion source, scaled down spacecraft models could be used, with the grid tied to the spacecraft ground, to look for possible arcing as the source is operated and study spacecraft charging.

Second, the probe data for the interior of the source indicated that the probe was very likely in the wall sheath when the electron and ion currents were nearly equal. These probe currents were nearly equal in the chamber also, when the plasma density was very low and the Debye length was 10cm. The conjecture then was that the volume contained by the wall sheath was not negligible compared to that of the chamber and the condition  $I_+ \approx I_-$  dominated. It would be interesting to study this effect by using a moveable probe system and an electric field measuring instrument to map out the fields close to and far from the walls, and close to the source where the plasma density is greater. This ion current effect should also be compared with that from ionospheric probe data. If  $I_+ \approx I_-$  in the ionosphere, it may suggest that there are electric fields in the ionosphere which are related to the earth's ionosphere acting as a sink for the plasma. The plasma in that region may in effect set up a wall type of sheath in which  $I_+ = I_-$  in order to prevent the rapid loss of electrons.

Third, the effects of large ion probe currents in a low

density plasma can be studied by stringing a series of grounded collectors along the chamber wall but insulated from it. Before biasing a probe or collector, each probe can be grounded through an electrometer to measure the wall current to that section of the chamber. Using a computer to switch from one wall probe to the other probe current data could be displayed on a CRT or printer. The process could be repeated when a probe or collector in the chamber is biased. In this way, the effect of a collector with a large ion current could be detected in the changing currents to the walls. These outputs can be used to map out the extent of the pre-sheath region and its dynamic growth as ion collector current increases.

#### REFERENCES

Langmuir, I. and Mott-Smith, H., *Phys. Rev.* 28 727-763 (1926)

Bohm, D., Burkap, E. H. S. and Massey, H. S. W., *The Characteristics of Electrical Discharges in Magnetic Fields*, Chapter 2. Edited by Guthrie, A. and Wokerling, R. K., McGraw Hill Inc. (1949)

Chen, F. F., *Introduction to Plasma Physics*, Chapter 1. Plenum Press (1974)

Chen, S. and Goodings, J. M., *Ion Current to a Langmuir Probe in Very Low Density Flowing Plasmas for Electron-Density Determinations*, *J. Appl. Phys.* 39 3300 (1968)

1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT

SPECTROSCOPIC STUDIES OF THYRATRON DISCHARGES

Prepared by:	Dr. Richard E. Miers
Academic Rank:	Associate Professor
Department and University:	Department of Physics Indiana U.-Purdue U/Fort Wayne
Research Location:	Wright Aeronautical Laboratories POOC-3
USAF research:	Dr. Alan Garscadden
Date:	September 21, 1984
Contract No:	F49620-82-003E

SPECTROSCOPIC STUDIES OF  
THYRATRON DISCHARGES

by

Richard E. Miers

ABSTRACT

Electron processes in hot and cold cathode thyratrons were studied using spectroscopic analysis of the discharges. The spectra emitted from a helium filled cold-cathode tube was studied as the tube was operated in a continuous discharge mode. Graphical results of intensity versus position between electrodes are presented for various He I and He II spectral lines. The time evolution of various spectral lines was measured for a specially designed ITT thyatron filled with a mixture consisting of 0.6 torr of hydrogen gas plus 10% helium gas. Photographic results are presented. Recommendations for follow-on research are presented.

### Acknowledgement

The author would like to thank the Air Force Office of Scientific Research and the Southeastern Center for Electrical Engineering Education for providing him with the opportunity to do research at Aero Propulsion Laboratory at Wright-Patterson AFB, Ohio. He wishes to acknowledge the Power Division, P00C-3 for their hospitality and for providing him with enjoyable working conditions. He would like to acknowledge all of the personnel of P00C-3 for their willing and able assistance in this project.

Finally he would like to thank Dr. Alan Garscadden for suggesting this research and for providing the guidance and equipment to carry it out.

## 1. INTRODUCTION:

The hydrogen discharge is of special interest because of its use as a high power plasma switch (thyatron), as a source of negative ions for the creation of fast neutral beams and as a source of free radicals for use in plasma chemistry and as a source of polarized atoms.

Much progress has been made in understanding the kinetics of energy transfer by treating the problem as an infinite medium and neglecting the influence of boundary conditions. This has been achieved usually by assuming an equilibrium behavior of the electrons and usually a Maxwellian or a Druyvesteyn distribution of electron energies.<sup>1,2</sup> However, all electrical discharges have to have a double-layer cathode sheath in order to satisfy the boundary condition of current continuity. In a hot-cathode discharge, this cathode sheath voltage is somewhat higher than the ionization potential of the gas. In a cold cathode discharge the cathode sheath voltage is around 300 volts. Thus, in either case, there is an injection of fast electrons into the discharge.<sup>3</sup>

The plasma that is used in the thyatron is then more characteristic of a negative glow at low currents, i.e. in the ignition state of the switch, than of a positive column discharge. As the plasma density increases, electron-electron collisions will cause efficient thermalization of the electron beam. The characteristics of the switch therefore change depending on pre-ionization, repetition-rate and the rate of rise of the trigger voltage.

There have been many experiments measuring the current-voltage characteristics of thyatrons.<sup>4,5,6</sup> However as indicated above, assumptions have to be made about the electron energy distribution in

order to estimate ionization and loss-rates. The scaling performance of such switches is not well defined.

The present project is concerned with a study of the electron energy distribution function of hot and cold cathode thyatron switches by spectroscopic analysis of discharges in selected gas mixtures. The addition of a small amount of Helium or Neon to a hydrogen discharge should not alter the ionization kinetics very much as their ionization potentials are much higher than those of either molecular or atomic hydrogen. The molecular hydrogen also is the only species discussed that has low energy inelastic rotational and vibrational energy losses, so the distribution function is controlled by the switch gas and the rare gases can be used as titration species.

The threshold for excitation of helium I spectral lines is considerably higher than the ionization potential of both atomic and molecular hydrogen. Also the helium II and neon II spectra can be used to identify excitation by high energy electrons.

The electron excitation functions and cross-sections for many of these lines have been measured.<sup>7 - 12</sup> There are also experimental or theoretical estimates for their transition probabilities, so that if quantitative measurements are made, then an estimate can be made of the actual number of fast electrons adjacent to the cathodes

## II. OBJECTIVES

The main objective of this project was to investigate the electron distribution functions in hot and cold cathode thyratrons through spectroscopic analysis of the discharges.

Three stages of the experiment were proposed. A commercial hot



cathode thyratron was used to establish the optical and electrical systems. A cold cathode gas filled triode designed by AFWAL/POOC-3 was studied. A specially modified ITT thyratron model that permits optical access to the cathode-grid space as well as to the grid-anode space was studied.

The data obtained from this research study will be used to benchmark computer models of pulsed power switches and to extend the productive capability of the codes.

### III. EXPERIMENTAL APPARATUS

The basic experimental arrangement is pictured in Figure 1 below.

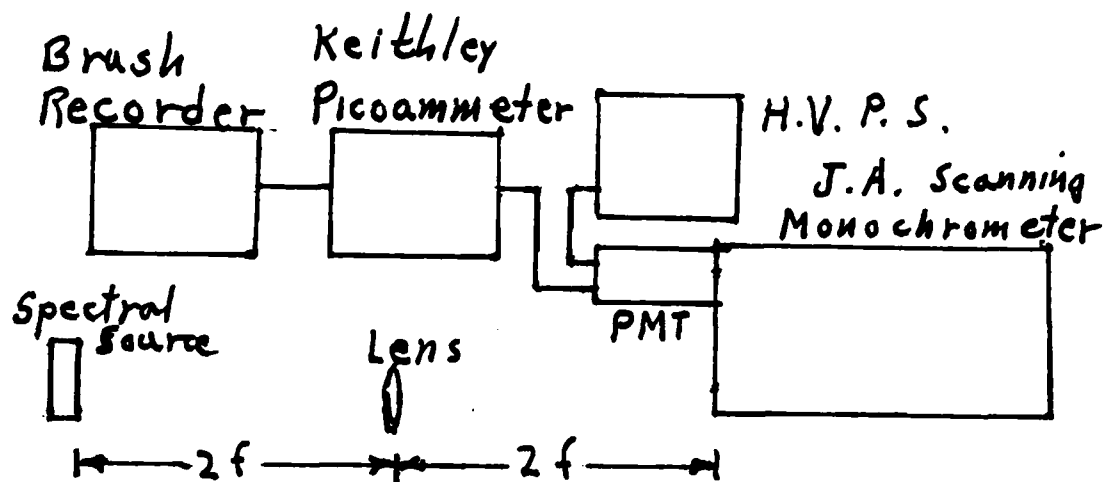


FIG 1 - BASIC EXPERIMENTAL ARRANGEMENT

The spectroscopic light source was placed at 300mm or  $2f$  from a lens with focal length  $f$  of 150mm. The variable slits of the Jarrell Ash 1/2 m scanning monochromator were located at a distance of  $2f$  from the lens giving unity magnification of the light source at the entrance

to the monochromator. The light signals were measured using an EMI 9659QB photomultiplier. To measure time averaged signals, the photomultiplier output was connected to a Keithly 440 digital picoammeter. The output of the picoammeter was fed into a Brush 550 recorder. For measuring the temporal response of fast light pulses the photomultiplier current output was connected in series with a 1000 ohm resistor. The voltage drop across the resistor was measured with a Tektronix digital oscilloscope. These voltage pulses were then recorded with an attached polaroid camera.

#### IV. EXPERIMENT

A helium-neon discharge tube was used to calibrate the wavelength drive of the monochromator and to measure the resolution of the monochromator at various slit width openings.

A commercial glass-enveloped thyatron was connected as shown in Figure 2.

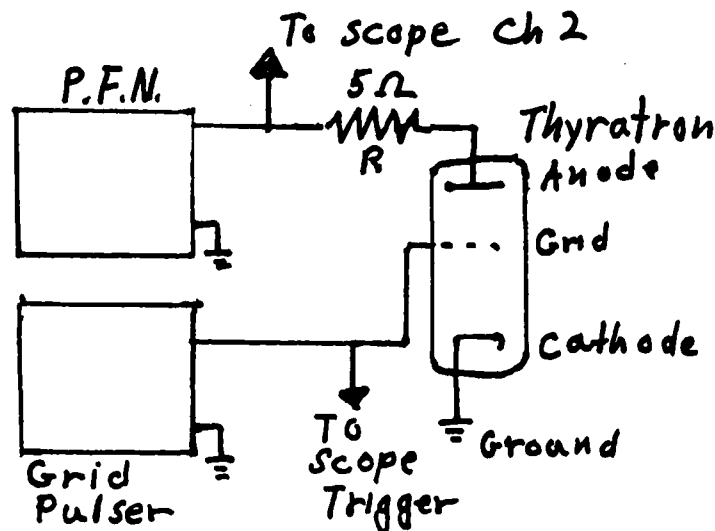


FIG 2 - BLOCK DIAGRAM OF THYRATRON DRIVE CIRCUIT

The pulse forming network (PFN) was designed to have a 5 ohm output

impedance and 10 microsecond pulse width when attached as shown to a triggered thyatron. The light emitted from the anode-grid space could be monitored by the optical system. By measuring the voltage drop across the 5 ohm current-limiting resistor in series with the thyatron one could determine the shape and magnitude of the current pulse output of the thyatron.

Figure 3 is a basic circuit diagram of the pulse forming network.

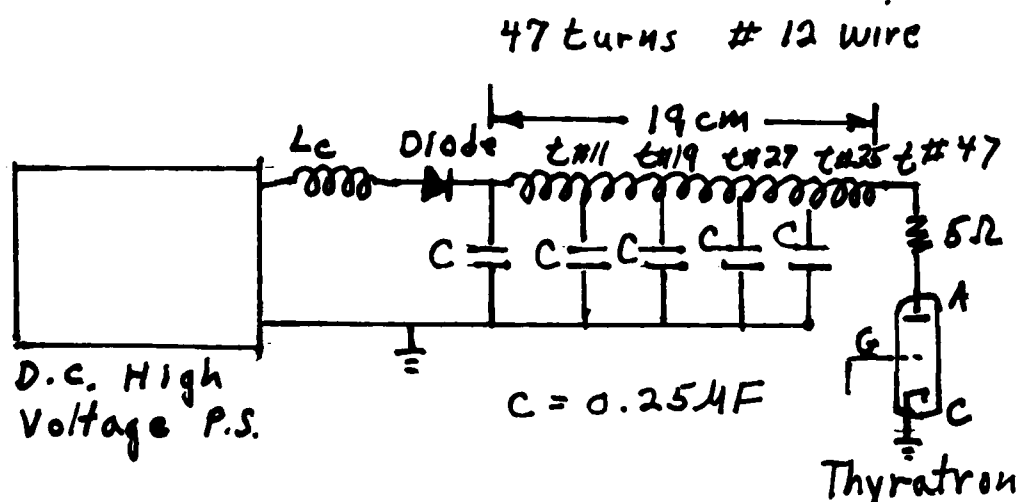


FIG-3- PULSE FORMING NETWORK CIRCUIT DIAGRAM

A tungsten ribbon standard lamp operated at 35 A was used to calibrate the optical system in the range of 3000 angstroms to 7000 angstroms.

A cold cathode tube was designed and constructed by P00C-3. A diagram of this tube and mount are shown in figure 4.

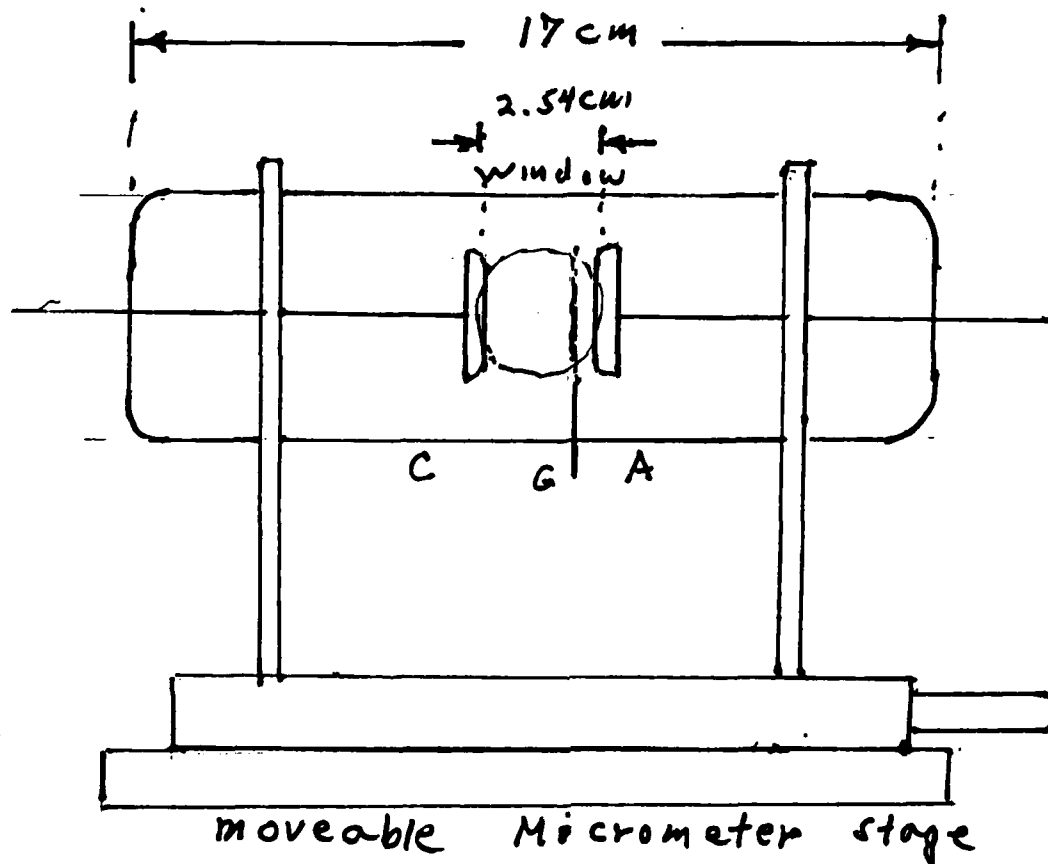


FIG-4 - DIAGRAM OF COLD CATHODE TUBE AND MOUNT

This tube was filled with 1.0 torr of pure helium. The tube was designed with a flat polished aluminum cold cathode and a flat aluminum anode about 2.54 cm in diameter. These electrodes were spaced about 2.36 cm apart with a grid composed of a metal ring with parallel wires spot welded in place located about 0.24 cm from the anode. An attempt was made to pulse this tube with the PFN and grid pulser. However, the tube would only operate in a continuous discharge mode. Therefore the grid was

allowed to float and the tube was operated with a continuous discharge using a potential difference between cathode and anode of around 300 V. The tube was mounted on a micrometer stage with a travel of slightly more than 2.54 cm allowing the spectrometer to scan the entire distance between cathode and anode. A current limiting resistor of 2,000 ohms was connected in series with the tube and power supply. The potential across this resistor was measured using a Simpson voltmeter. This potential could be varied from 0.4 V. to 1.5 V indicating currents in the discharge of from 0.20 mA to 0.75 mA. The optical system was aligned with a helium-neon laser to center the image of the window on the entrance slits of the monochromator.

The image of the center of the anode grid space was focussed onto the monochromator slit. The relative intensity of the helium II ion line at 4586 angstroms was measured as a function of discharge current. Within the uncertainty of the measurements this intensity was linear with the current. The same check was made for the helium I atomic spectral line at 4713 angstroms with the result that the intensity was also linear with the discharge current.

The relative spectral intensities of five helium I lines and two helium II lines were measured as functions of position between the anode and cathode. The results are graphed in figures 5, 6 and 7.

In these graphs the image of the cathode position is located at approximately 0.2 cm with the image of the anode at approximately 2.5 cm. The dip in intensity at 2.0 cm shown in each of the graphs is due to image of the grid located at this position.

Each of the three triplet lines is more sharply peaked near the center of the tube. The two singlet lines show a relatively higher

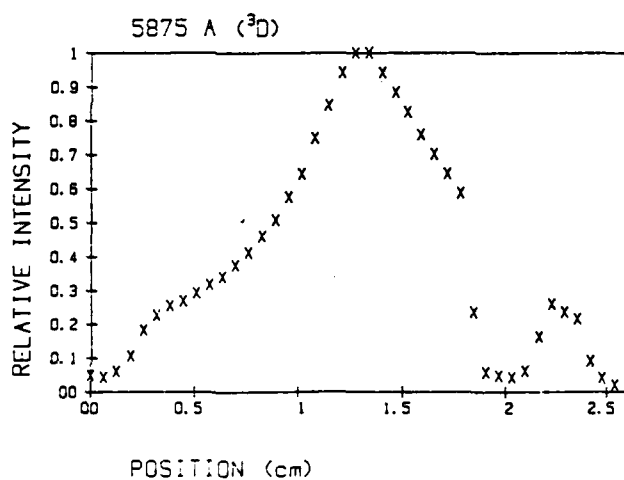
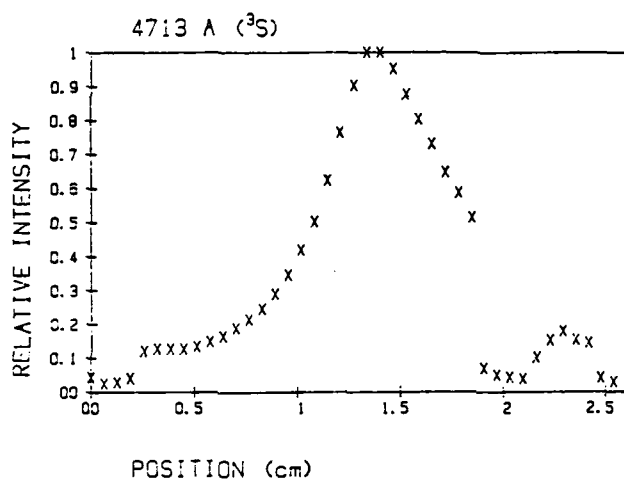
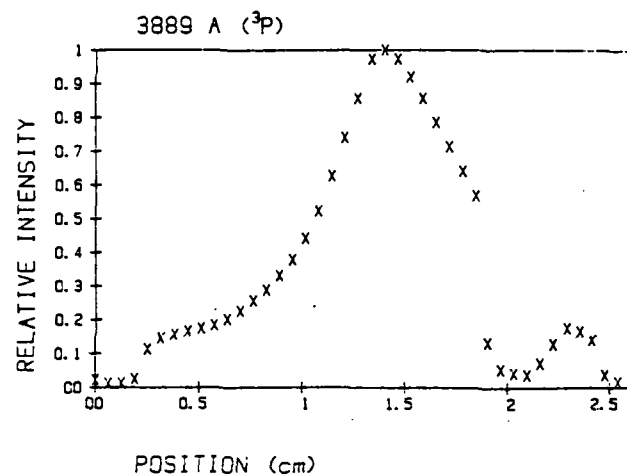


FIG 5 - Relative Intensity of Triplet He Lines vs. Position

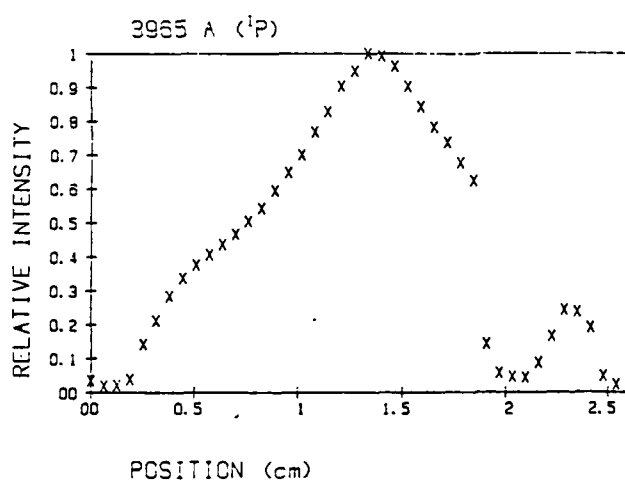
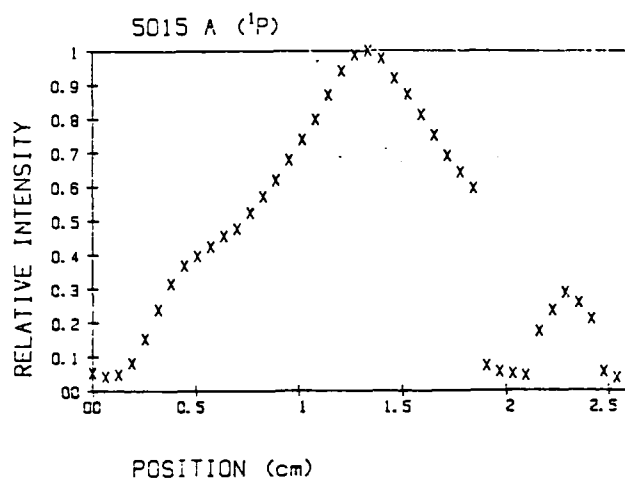


FIG 6 - Relative Intensity of Singlet He Lines vs. Position

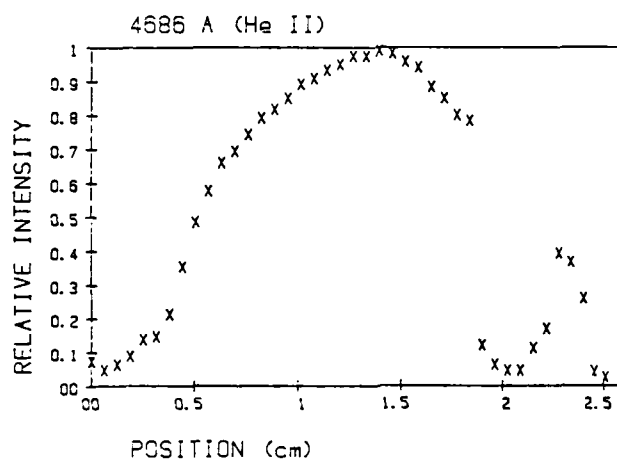
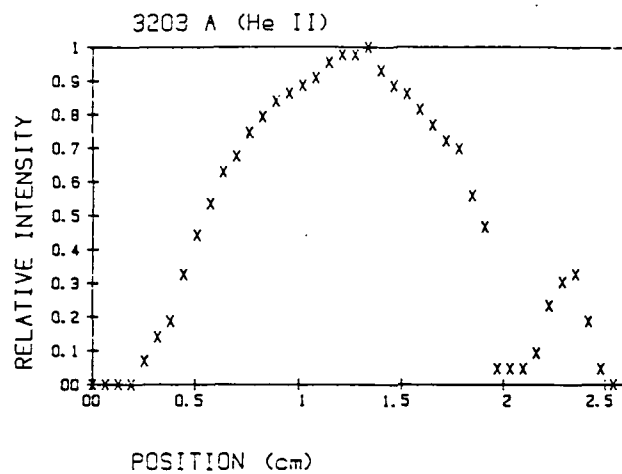


FIG 7 - Relative Intensity of He Ion Lines vs. Position



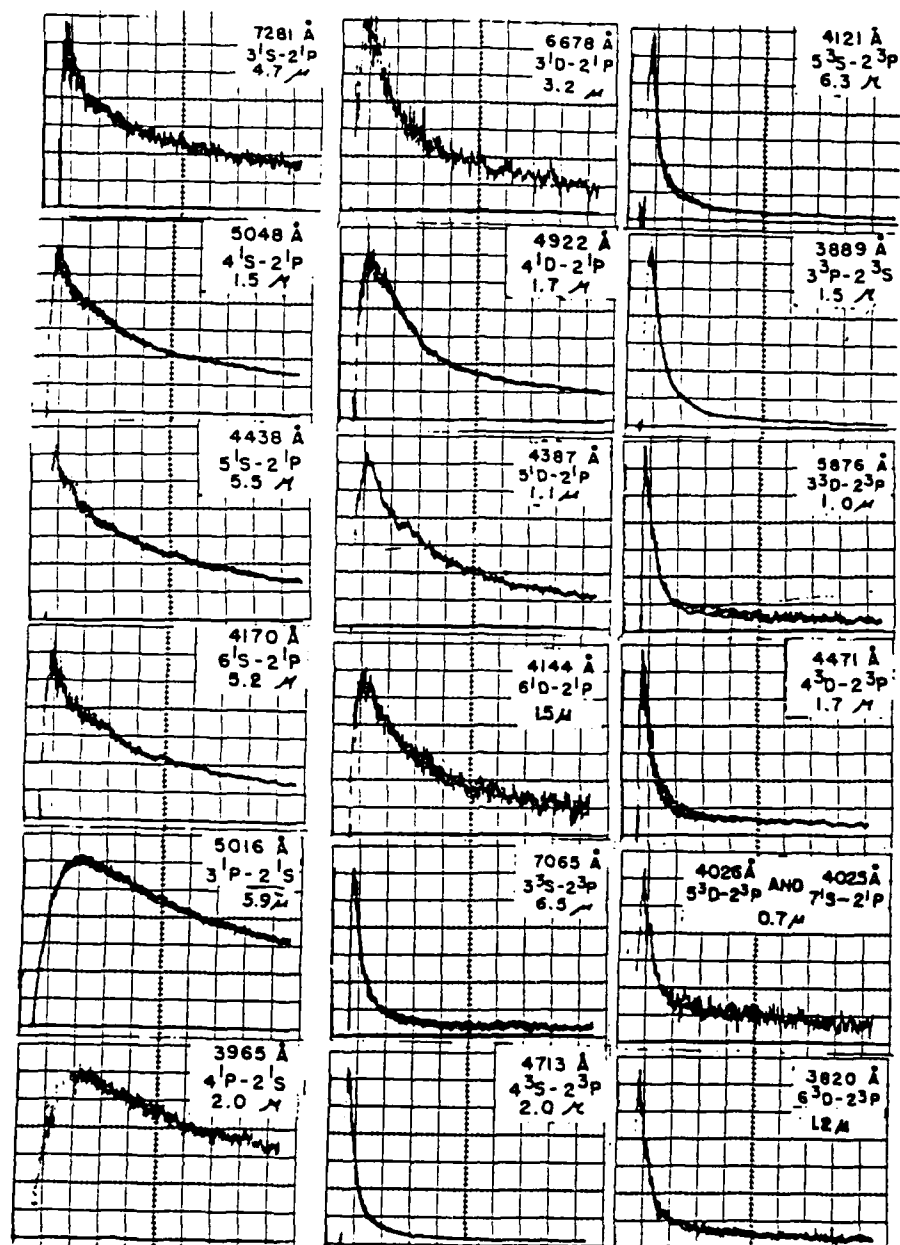


FIG 8 - Apparent Excitation Functions of Helium Lines

From St. John, et al. Reference 10. (Electron energy is from 0 to 500 eV)

onset of intensity near the cathode. If one examines the apparent excitation functions for these lines as measured by St. John et al.,<sup>10</sup> (illustrated in figure 8) this would seem to indicate a number of higher energy electrons near the cathode.

The two visible He II lines show a rapid rise in intensity at positions much closer to the grid indicating a number of electrons with energies of around 75 to 200 eV, which represent the onset and peak of the electron excitation functions of these lines.<sup>7</sup>

The specially modified ITT thyratron was connected as shown in Figure 2. The tube was filled with a mixture consisting of 0.6 torr of hydrogen gas and 0.06 torr of helium. The time evolution of the light pulses at the various spectral wavelengths of hydrogen and helium were measured using a Tektronix 7904 oscilloscope and a Tektronix C-53 camera. The light pulses were measured with the photomultiplier output connected in series with a 1000 ohm resistor. The voltages across the resistor were measured with the oscilloscope and recorded with the camera. Figures 9 and 10 show the typical time evolution of the spectral lines of hydrogen and helium occurring midway between the anode and grid of the thyratron. Figures 11 and 12 show the typical time evolution of the spectral lines emitted from midway between the grid and cathode viewed through holes drilled in the grid structure.

#### V. RECOMMENDATIONS

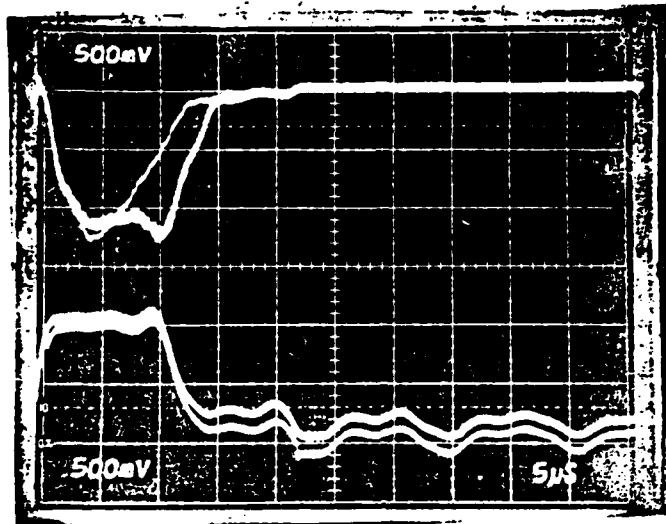
A theoretical analysis of these data will be undertaken with a view toward understanding processes taking place in thyratron tubes, particularly the nature of the electron distribution function. Hopefully, by computing a trial fit of various possible distribution

functions making use of the known cross-sections for excitation for the various spectral lines one can get some understanding of these functions.

One recommendation for follow-on research is to undertake this theoretical study of the distribution functions in the tubes.

A second recommendation for follow-on research is to continue study of the cold-cathode tube using the techniques described in this report. The tube will be filled with various mixtures of gases. Much of the equipment needed to begin this research is already available to me at my home institution.

Hydrogen 6563 A



Hydrogen 4340 A

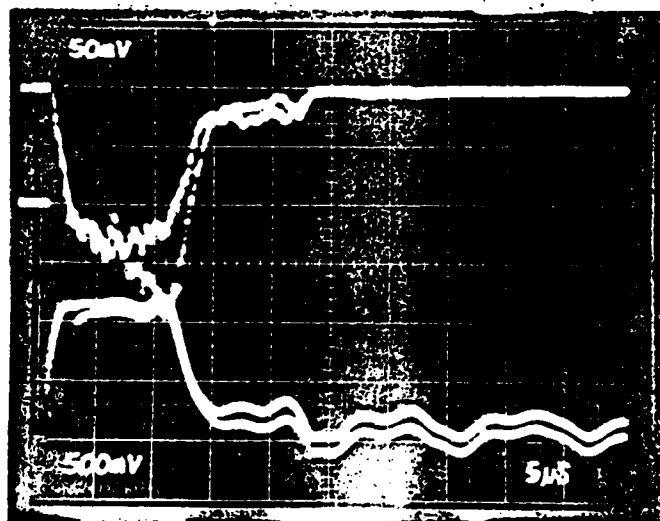
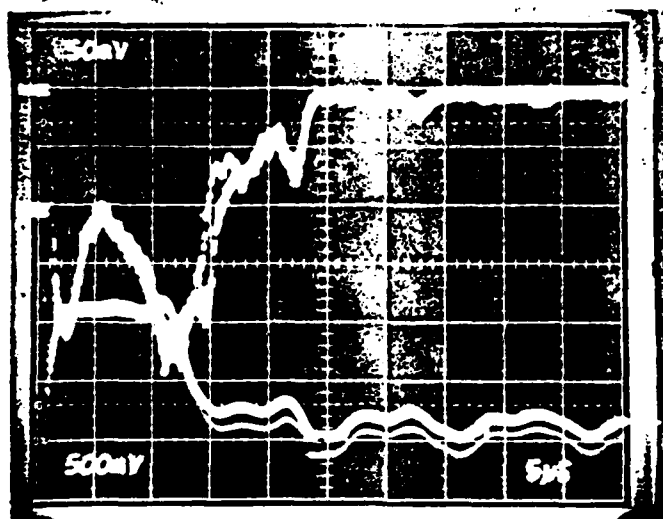


FIG 9 - Typical Time Evolution of H Atomic Spectral Line Between Anode and Grid of the Thyatron  
The upper trace in each photo is the voltage across the 1000 ohm resistor in series with the photomultiplier.  
The lower trace in each photo is the voltage output of the PFN/1000 and is proportional to the thyatron current pulse.

Helium 3889 A



Helium 5876 A

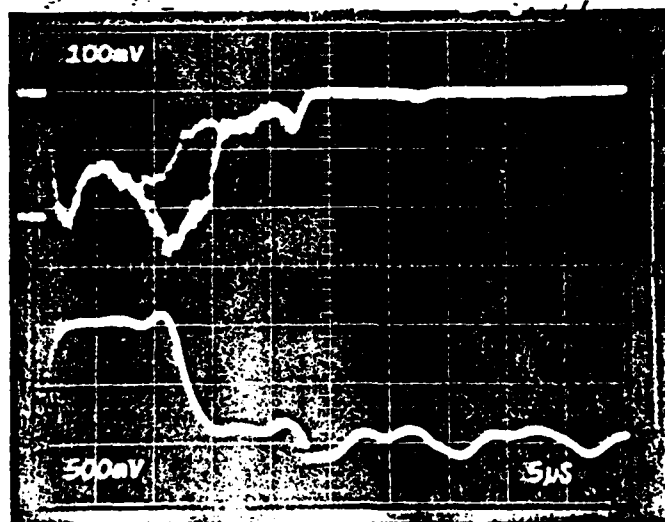


Fig 10 - Typical Time Evolution of He Atomic Spectral Line Between Anode and Grid of the Thyatron

Hydrogen 6563 A



Hydrogen 4340 A

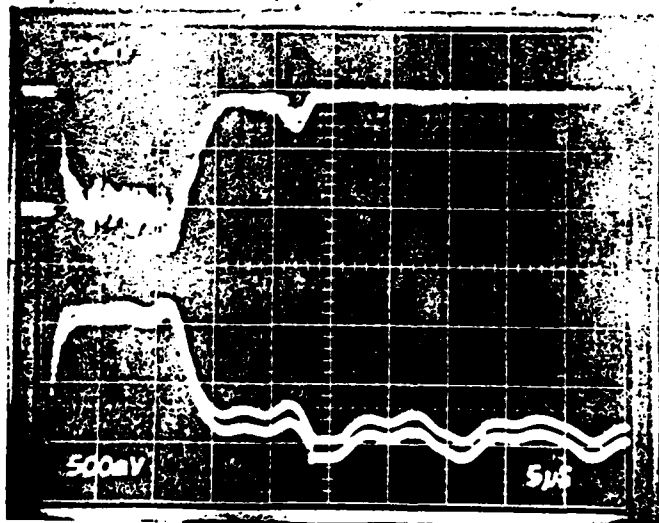
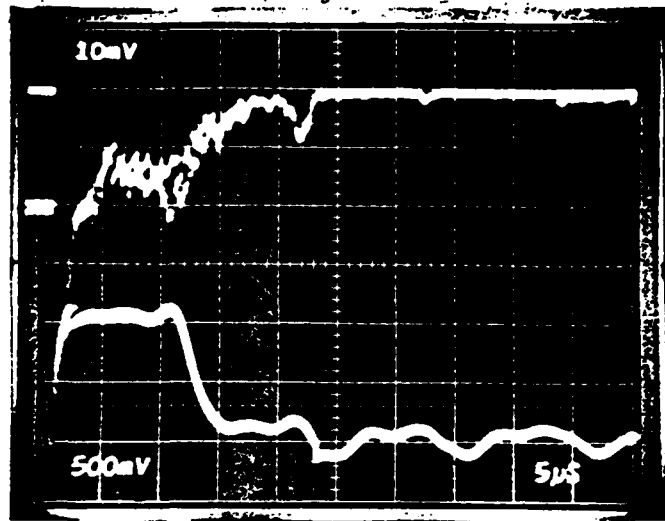


FIG 11 - Typical Time Evolution of H Atomic Spectral Line  
Between Cathode and Grid of the Thyatron

Helium 3889 A



Helium 5876 A

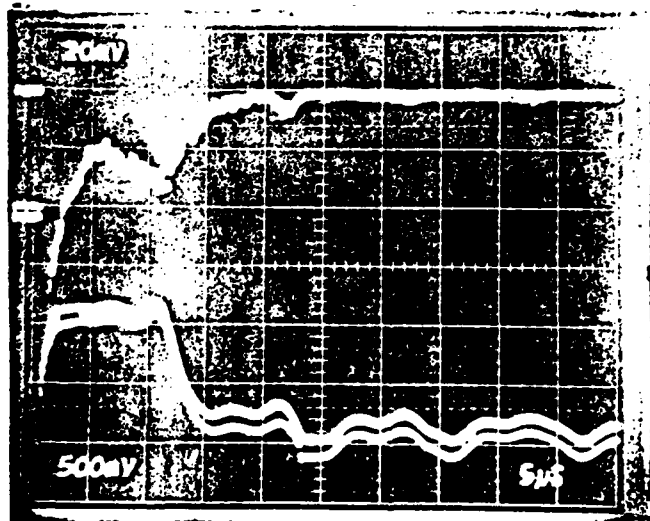


FIG 12 - Typical Time Evolution of He Atomic Spectral Line Between Cathode and Grid of the Thyatron

#### REFERENCES

1. J. A. Kunc, S. Guha and M. A. Gunderson, "A Fundamental theory of high power thyratrons I: The electron temperature," Lasers and Particle beams (1983), vol 1 pp 395 - 405.
2. J. A Kunc and M. A. Gunderson, "A fundamental theory of high power thyratrons II: The production of atomic hydrogen and positive ions," Laser and Particle beams (1983), vol 1 pp 407 - 425.
3. A. Garscadden, "Electron kinetics as applied to Pulsed Power Switching," 4th IEEE Pulsed Power Conference, Albuquerque, N. M., (1983) pp 455 - 461.
4. J. Creedon and S. Schneider, "Megawatt Average Power Adiabatic Mode Thyatron," Proc. of the Int. Pulsed Power Conf., Lubbock, Texas, (1976) pp 184-1 - 184-6.
5. S. Goldberg and J. Rothstein, "Hydrogen thyatron," in Advances in Electronics and Electron Physics Vol XIV, Acad. Press, pp 207 - 264 (1961).
6. S. Friedman, S. Goldberg, J. Hamilton, S. Merz, R. Plants and D. Turnquist, "Multi-gigawatt Hydrogen Thyatron with nanosecond rise times," Private Communication.
7. R. M. St. John and C. C. Lin, "Production of Excitation and Ionization in Helium by Single-Electron Impact," J. Chem. Phys., Vol 41, (1964) pp 195 - 197
8. H. R. Moustafa Moussa and F. J. De Heer, "Excitation of Helium to Excited Ion States by Electron and Proton Impact," Physica, Vol 36, (1969) pp 517 - 549.
9. K. G. Walker and R. M. St. John, "Simultaneous Ionization and Excitation of Neon by Electron Impact," Physical Review A, Vol 6, (1972) pp 240 - 250.
10. R. M. St. John, F. L. Miller, and C. C. Lin, "Absolute Electron Excitation Cross Sections of Helium," Physical Review, Vol 134, (1964) pp A888 - A879.
11. H. R. Moustafa Moussa, F. J. De Heer and J. Schutten, "Excitation of Helium by 0.05 - 6 keV Electrons and Polarization of the Resulting Radiation," Physica, Vol 40, (1969) pp 517 - 549.
12. J. G. Showalter and R. B. King, "Absolute Measurement of Total Electron-Impact Cross Sections to Singlet and Triplet Levels in Helium," Physical Review A, Vol 11, (1975) pp 1899 - 1910.



1984 USAF-SCEEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the  
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the  
SOUTHEASTERN CENTER FOR ELECTRICAL ENGINEERING EDUCATION

FINAL REPORT  
ON INTERFACING LOGIC PROGRAMMING  
SYSTEMS AND RELATIONAL DATABASES

Prepared by: John T. Minor  
Adademic Rank: Assistant Professor  
Department and University: School of Electrical Engineering and  
Computer Science, The University of Oklahoma  
Research Location: Rome Air Development Center, Command and  
Control Division, Software Technology  
Branch, Knowledge Engineering Section  
USAF Research Contact: Robert Schrag and Dr. Northrop<sup>u</sup> Fowler  
Date: September 18, 1984  
Contract No: F49620-82-C-0035

ON INTERFACING LOGIC PROGRAMMING  
SYSTEMS AND RELATIONAL DATABASES

By

John T. Minor

ABSTRACT

Current logic programming systems are inadequate for the implementation of knowledge-based systems which require a large number of facts and also want quick real-time responses to queries. To improve this situation, the possibility of using a relational database system as a secondary memory to hold and retrieve facts for the logic programming system was investigated. An interface between an actual logic programming system (LM-Prolog running on a LISP machine) and a database machine (IDM-500) was implemented, using the compiled-procedure approach. This approach was found to be the best for knowledge-based systems which have unchanging, moderately-sized intentional databases. A procedure for handling recursive rules in this approach was developed by the author and is described. Further problems which were discovered in the development of this interface, such as the handling of functional terms or side-effect operators, are presented and offered as suggestions for further research.

### ACKNOWLEDGEMENTS

The author would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and The Southeastern Center for Electrical Engineering Education for providing him with the opportunity to spend a very worthwhile summer at Rome Air Development Center, Griffiss AFB, N.Y. I would like to acknowledge the laboratory, in particular, Sam DiNitto of the Software Technology branch and Don Roberts of the Knowledge Engineering section, for their hospitality.

Finally, I would like to thank our Tuesday morning discussion group for their inspiration and guidance: Bob Schrag, Dr. Nort Fowler, Dr. Gerry Capraro, Dr. Ray Liuzzi, Dr. Bruce Berra of Syracuse University, and Nancy Kirkwood of Colorado State University.

## I. INTRODUCTION:

A logic programming system (LPS) stores a set of concrete facts and logical rules in its local memory, and then handles queries through a theorem-prover which returns one proof (answer) at a time. Using this design for implementation of a knowledge-based system (KBS) has several drawbacks:

- (1) A KBS has an extremely large set of facts to be stored, which will cause immense paging and searching problems in the LPS.
- (2) A KBS usually wants all results to a query, while a LPS is designed for one-answer-at-a-time retrieval.
- (3) The real-time response to a query will be slow because of (1) and (2), and because of the fact that theorem-proving (a combination of deductive search, unification, renaming and substitution operations) must be performed on each query to the system.

An approach which has been proposed to improve this situation is shown in Figure 1.<sup>1,2,3</sup> Here, a database management system (DBMS) is added to hold the large set of facts, or extensional data (EDB). Only the set of rules, or intentional data (IDB), is stored in the memory of the LPS. It is believed this design will have a much quicker response time to queries for the following reasons:

- (1) The DBMS is designed for efficient storage and searching of large sets of facts.

- (2) The DBMS retrieves tables of answers at a time.
- (3) The IDB can be kept in a modified form where little or no theorem-proving is done at query response time.

The DBMS that is used is normally of the relational model because of the natural correspondence between a tuple in a relation and a fact (unit ground clause) in logic. For this reason we will assume in the following that the LPS is accessing a relational database (RDB). Other models and other ways of handling secondary storage will not be considered in this paper.

The architecture of figure 1 raises a number of questions, i.e. how should the interface ( $\longleftrightarrow$ ) be implemented, what is the most efficient form in which to keep the logical rules, and what problems will be caused by connecting these two systems which have very different design motivations. These are the issues addressed in this paper.

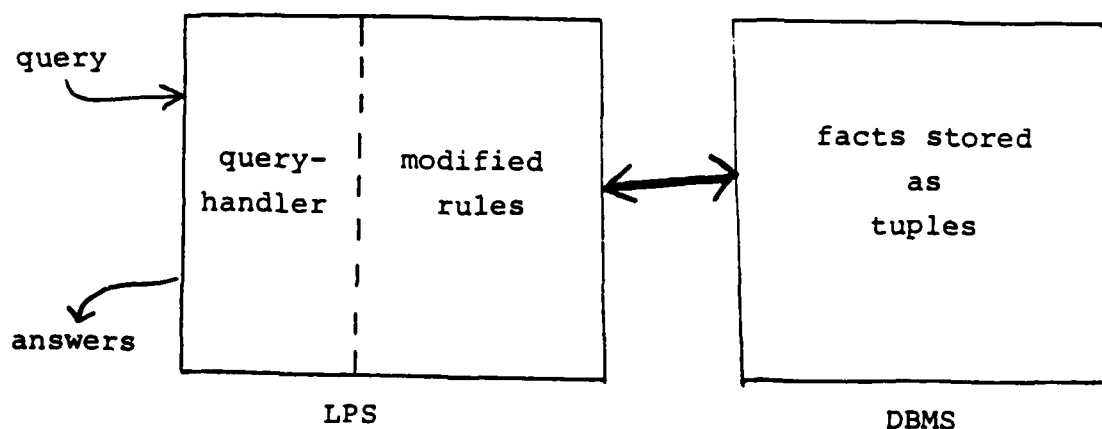


Figure 1. A Proposed KBS Architecture

## II. OBJECTIVES:

The objective of this research was to look at techniques for interfacing logic programming systems and relational databases in order to provide support for knowledge-based and other artificial intelligence work at the Rome Air Development Center. Recommendations on the choice of interfacing approach are given in the last section of this paper. Also investigated were problems caused by differences in the capabilities of logic programming systems and relational databases. In previous research, these conflicts were "solved" by putting large restrictions on the LPS, e.g. no functional terms, no recursion, etc. These systems were actually "databases with deductive facilities" in Gallaire's<sup>4</sup> categorization. The philosophy of this work is rather to look at the problems of developing a "logic programming system with database facilities."

The different interface techniques and problems were investigated, and some solutions tested, by implementing an interface between an actual LPS and actual RDB at the Rome Air Development Center. Specifically, LM-Prolog<sup>16</sup> running on a LMI Lambda machine was used as the LPS, and a IDM-500, a backend database machine, was used as the RDB.

AD-A154 337

UNITED STATES AIR FORCE SUMMER FACULTY RESEARCH PROGRAM 15/15  
(1984) PROGRAM MA. (U) SOUTHEASTERN CENTER FOR  
ELECTRICAL ENGINEERING EDUCATION INC S.

UNCLASSIFIED

W D PEELE ET AL. DEC 84 AFOSR-TR-85-0480

F/G 5/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A



### III. COMMUNICATION TECHNIQUES:

Much research has already addressed the problem of when a LPS should request data from the RDB while processing a query. Three distinct techniques have <sup>evolved</sup> solved:

(1) Interpreted approach.<sup>5,6</sup> This technique is a direct extension of the standard LPS control mechanism, in which whenever the theorem-prover needs to access a fact, it immediately sends a request to the RDB and waits for an answer before continuing. It thus has the advantages of being easy to implement and of recognizing dead-end deduction paths sooner than the compiled methods. On the other hand, because relations are called one at a time, no optimizations can be performed on the RDB requests and joins must be done inside the LPS.

(2) Compiled-query approach.<sup>7,8,9,10</sup> Deductive search continues as in the interpreted approach until a fact is needed. At this point, the fact is assumed to be true and the deductive search continues. When a solution is found, the assumed-true facts are sent as a join to the RDB to be verified or rejected. This intermediate object which is sent to the RDB is often called a proof plan, since it is not a complete solution until the RDB verifies it. Because an entire proof plan is constructed before it is sent to the RDB, this technique will, in general, make fewer calls to the RDB than the interpreted method, and will execute the joins in the RDB machine which is designed to do this

operator effectively. Also local optimization on a proof plan is now possible. Local optimization refers to the reordering of calls within a proof plan in order to speed-up the join operation <sup>11</sup>.

(3) Compiled-procedure approach.<sup>12,13</sup> In a once-only step before query processing, the theorem-prover is repeatedly applied to all rules in the IDB until they are reduced to a set of proof plans. Then the original IDB is replaced with this modified one. A query can now be processed by simply collecting its proof plans and sending them to the RDB for verification or rejection. This approach has the great advantage of completely eliminating theorem-proving from the response time to a query, since theorem-proving needs to be done only in the pre-processing step. Local and global optimization (the recognition of shared structures between proof plans)<sup>12</sup> can be done very easily since all proof plans exist before any call to the RDB is made. Teiter<sup>13</sup> has shown that any non-recursive IDB can always be compiled in this manner, and Section V shows how this approach can be extended to recursive rules.

Examples which illustrate the differences between approaches (1) and (2) can be found in Chakravarthy, et al.<sup>5</sup> Approach (3) is discussed in more detail in the following sections.

#### IV. AN IMPLEMENTED INTERFACE:

As part of this research, a partial interface between LM-Prolog and IDL<sup>17</sup> the input language to the IDM-500 database machine was completed. The interface program was implemented in LISP to run on a LISP machine, and the compiled-procedure approach was used, for reasons listed in Section VI. The overall design is shown in Figure 2.

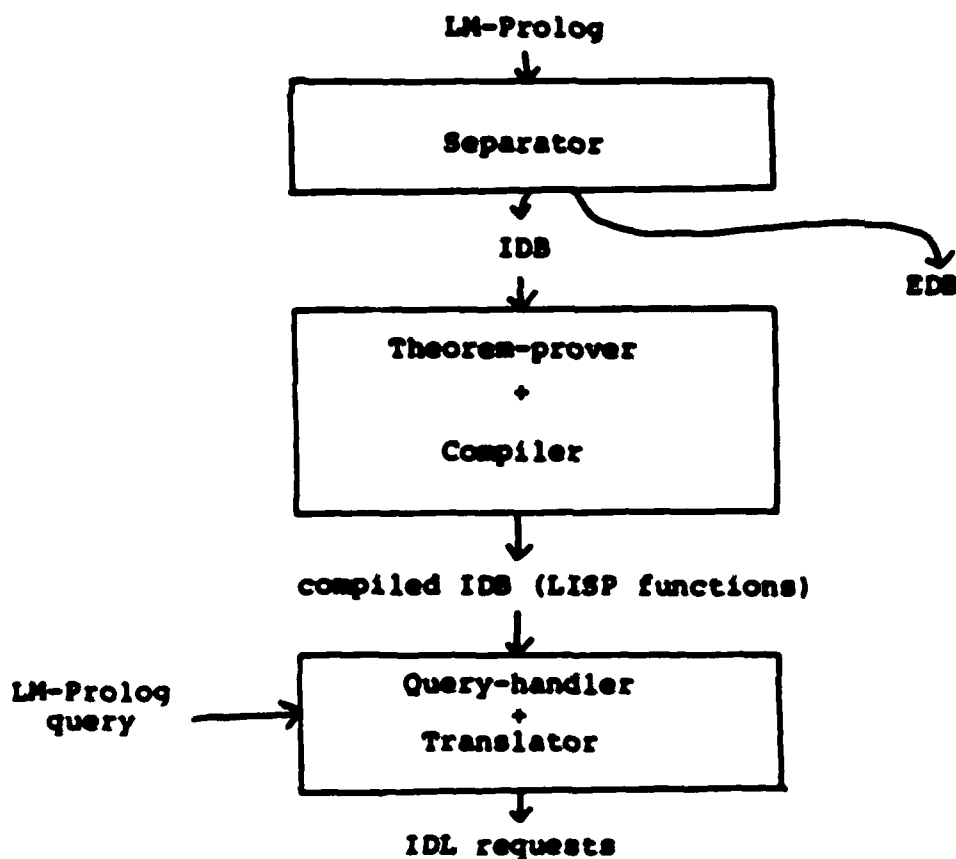


Figure 2. Design of Interface

The separator pulls out any facts from the input and sends them to the RDB for storage. Also if a predicate name

appears in both the EDB and IDB, the separator rewrites it using the following procedure: Assume both  $P(X)$ -body. and  $P(a)$ . appear in the input file, then a new predicate name  $P'$  is generated, and the above rules are rewritten as  $P(X)$ :-body.  $P(X)$ :-  $P'(X)$ . and  $P'(a)$ . This logically equivalent form completely separates IDB names,  $P$ , from EDB names,  $P'$ .

Using the compiled-procedure approach, the theorem-prover-compiler then translates each LM-Prolog IDB procedure into a LISP function whose body is simply the disjunction of all possible proof plans for that predicate. This procedure is of course impossible for recursive predicates, so a slightly different form of LISP function is generated for them. The technique for recursive procedures is described in the next section. Note that LISP was selected as the internal form of the compiled predicates because this work was done on a LISP machine.

At this point the original IDB and the theorem-prover can be expunged from the system. Any query to predicate  $Q$  can now be handled as a call to the generated LISP function  $Q$  in the compiled IDB. The query-handler-translator then converts the proof plans returned into disjunctive normal form and translates them into IDL. For recursive predicates, the query-handler also controls the maximum number of proof plans generated.

As an example, consider the LM-Prolog input file shown in Figure 3. During the separation step, the "father" and

The form of "ancestor," a recursive predicate, will be discussed in the next section.

#### V. RECURSIVE RULES:

A predicate defined by a recursive rule will have an infinite number of possible proof plans associated with it. Therefore it is impossible to explicitly represent its compiled form as a finite set of proof plans. Instead an implicit form must be used. The following form and technique were developed by the author, and is presented here without proof.

Assume  $P$  is a direct or indirect recursive predicate which does not use multiple recursion i.e.,  $P$  does not call itself more than once directly or indirectly, and does not use mutual recursion i.e.,  $P$  does not call another recursive predicate directly or indirectly. Then by applying the theorem-prover, we can always rewrite the rules for  $P$  in the following form:

$$\begin{aligned} P(\bar{X}) &:- \mathcal{E}_1(\bar{X}). \\ &\vdots \\ P(\bar{X}) &:- \mathcal{E}_k(\bar{X}). \\ P(\bar{X}) &:- \mathcal{A}_1(\bar{X}, \bar{Y}_1, \bar{Z}), P(\bar{Z}). \\ &\vdots \\ P(\bar{X}) &:- \mathcal{A}_m(\bar{X}, \bar{Y}_m, \bar{Z}), P(\bar{Z}). \end{aligned}$$

where  $\mathcal{E}_1$  and  $\mathcal{A}_1$  are expressions formed entirely of EDB predicates and  $\bar{X}$ ,  $\bar{Y}_1$ , and  $\bar{Z}$  are vectors of variables. The

arguments in all the heads of the rules have been made the same,  $\bar{X}$ , by renaming variables and by replacing non-variable arguments using the following technique: Assume  $t_1$  is a non-variable term in rule  $P(\dots, t_1, \dots) :- \text{body}$ . Then rewrite the rule as  $P(\dots, X_1, \dots) :- X_1 = t_1, \text{body}$ . This same procedure is used to make the arguments in all the recursive calls equal to  $\bar{X}$ . All other variables in the  $Q_1$  expression i.e., not in  $\bar{X}$  or  $\bar{Z}$ , are collected together into the set  $\bar{Y}_1$ . Non- $\bar{X}$  variables in  $\bar{E}_1$  are not shown but are assumed to be distinct.

Now the above set of rules can be replaced by the equivalent LISP function:

```
(DEFUN P( $\bar{X}$  K)
  (COND((EQ K 0) *(( $\bar{E}_1(\bar{X})$  OR ... OR  $\bar{E}_K(\bar{X})$ )))
  (T(LET(( $\bar{Z}$ (GENVECTOR))( $\bar{Y}_1$ (GENVECTOR))...( $\bar{Y}_m$ (GENVECTOR)))
    (APPEND(P  $\bar{Z}$  (SUB1 K))
      *(AND ( $Q_1(\bar{X} \bar{Y}_1 \bar{Z})$  OR ... OR  $Q_m(\bar{X} \bar{Y}_m \bar{Z})$ )
        )))))
```

where GENVECTOR generates new symbols for all the variables in the vector, and \* is a modified form of quote which returns its argument unevaluated except that  $\bar{X}$ ,  $\bar{Y}$  and  $\bar{Z}$  variables are instantiated. On a call to P with  $K = n$ , the proof plans for n executions of the recursive loop will be returned in conjunctive normal form.

The query-handler controls the execution of P by calling P with K = 0, 1, ..., n, then converts the result of each call into disjunctive normal form and translates the resulting proof plans into IDL. The value of n at which calls to P are terminated is, at this time, set by the user. This is an inadequate solution, but more work is needed on this problem before a complete set of termination criteria is found. Henschen<sup>14</sup> has given some ideas on this topic.

As an example of this technique, look at the "ancestor" definition in Figure 3. After applying the theorem-prover, we can rewrite this definition as the set of rules:

```
((ancestor ?x1 ?x2)(father ?x1 ?x2))
((ancestor ?x1 ?x2)(mother ?x1 ?x2))
((ancestor ?x1 ?x2)(father ?y1 ?x2)( = ?z1 ?x1)( = ?z2 ?y1)
                                     (ancestor ?z1 ?z2))
((ancestor ?x1 ?x2)(mother ?y1 ?x2)( = ?z1 ?x1)( = ?z2 ?y1)
                                     (ancestor ?z1 ?z2))
```

These rules can then be replaced by the LISP function:

```
(DEFUN ANCESTOR(X1 X2 K)
  (COND ((EQ K 0)(LIST (LIST (LIST 'FATHER X1 X2))
                           'OR (LIST (LIST 'MOTHER X1 X2))))
        (T (LET ((Z1 (GENSYM))(Z2 (GENSYM))(Y1 (GENSYM)))
              (APPEND (ANCESTOR Z1 Z2 (SUB1 K))(LIST 'AND
```

```

(LIST (LIST (LIST 'FATHER Y1 X2)(LIST '= Z1 X1)
            (LIST '= Z2 Y1)))
'OR (LIST (LIST 'MOTHER Y1 X2)(LIST '= Z1 X1)
      (LIST '= Z2 Y1))
))))))

```

Calls to ANCESTOR as well as to the compiled forms of "parent," "grandparent," etc. are shown in Figure 4. The output you see are the proof plans which have been translated into IDL.

A multiple recursive rule, such as  $P(\bar{X}) :- Q(\bar{X}, \bar{Y}, \bar{W}, \bar{Z}), P(\bar{W}), P(\bar{Z})$ , can be treated as the single recursive rule  $P(X) :- Q'(\bar{X}, \bar{Y}', \bar{Z}), P(\bar{Z})$ . If we let  $Q' = Q, P(\bar{W})$  and let  $\bar{Y}' = \bar{Y} \cup \bar{W}$ . The previously described technique can now be used to generate a recursive function for this  $P$  with the modification that the  $P(\bar{W})$  inside  $Q'$ , since it is not an EDB predicate, will have to be treated as another recursive call to  $P$ . Mutual recursive rules like  $P(\bar{X}) :- Q(\bar{X}, \bar{Y}, \bar{W}, \bar{Z}), Q(\bar{W}), P(\bar{Z})$ , where  $Q$  is recursive, can be handled in the same manner. Note that these two recursive cases were not handled in the implemented interface, but they are a straight-forward extension.

## VI. RECOMMENDATIONS:

If the KBS has a moderately-sized, seldom-changed IDB then the compiled-procedure approach becomes the best interface technique for the following reasons:



```

(define-predicate grandparent
  ((grandparent ?x ?y)(parent ?x ?z)(parent ?z ?y)))
(define-predicate parent
  ((parent ?x ?y)(mother ?x ?y))
  ((parent ?x ?y)(father ?x ?y)))
(define-predicate father
  ((father fred mary))
  ((father george james))
  ((father john fred)))
(define-predicate mother
  ((mother sue mary))
  ((mother jane sue))
  ((mother liz fred))
  ((mother sue james)))
(define-predicate ancestor
  ((ancestor ?x ?y)(parent ?z ?y)(ancestor ?x ?z))
  ((ancestor ?x ?y)(parent ?x ?y)))

```

Figure 3 - Example LM-Prolog Program

"mother" definitions will be removed and sent to the RDB for storage. Then the LDB predicates will be compiled into LISP functions. The "parent" LISP function will look like

```

(DEFUN PARENT(X Y)
  (LIST (LIST (LIST 'FATHER X Y))
        'OR (LIST (LIST 'MOTHER X Y)))))

```

and "grandparent," which has four proof plans, will look like

```

(DEFUN GRANDPARENT(X Y)
  (LIST (LIST (LIST 'FATHER X ?Z)(LIST 'FATHER ?Z Y))
        'OR (LIST (LIST 'FATHER X ?Z)(LIST 'MOTHER ?Z Y))
        'OR (LIST (LIST 'MOTHER X ?Z)(LIST 'FATHER ?Z Y))
        'OR (LIST (LIST 'MOTHER X ?Z)(LIST 'MOTHER ?Z Y)))))

```



**Figure 4. (continued)**

**Figure 4. (continued)**

- (1) Query response time is greatly enhanced because no deductive search or theorem-proving is conducted.
- (2) There is no need for a theorem-prover at query time, so it can be expunged from the system after compilation along with the original IDB.
- (3) Joins of tables are done inside the RDB instead of in the LPS.
- (4) It has the capability to perform local and global optimizations of proof plans easily.
- (5) The overhead of compilation can be ignored because the IDB is seldom modified.
- (6) The increase in size of the compiled IDB over the original IDB will not cause complications because of the moderate size of the original IDB.

In applications where the IDB is large or is being modified frequently, the compilation overhead and the increased size of the compiled IDB may become significant problems in the compiled-procedure approach. In those cases the compiled-query approach with the IDB kept as a predicate connective graph will have an advantage. The predicate connective graph<sup>15</sup> form reduces theorem-proving at response time, while not increasing the size of the IDB. In this representation when a rule is entered into the system, arcs are set up from the IDB predicates in the body of the rule to the heads of rules defining that predicate. Each arc is then labeled with the most general unifier of the two literals which it connects. This pre-processing of the IDB

will eliminate the deductive search and part of the unification that would normally be done at query response time in the compiled-query approach. See Kellogg et al.<sup>8</sup> for an example of this approach.

#### Further Work:

A number of simplifying assumptions were made in the building of the LM-Prolog to IDL interface. Some straightforward enhancements that could be made to the interface next would be to implement meta-predicates and variable-arity predicates as allowed in LM-Prolog, implement multiple and mutual recursion as described in Section V, and handle multiple-literal queries.

More difficult issues, which will require investigation, include:

- (1) How side-effect operators like assert, print, cut, etc. should be incorporated into the compiled schemes.
- (2) How and when to do local and global optimization in the interface.
- (3) How to handle functional-terms and list-terms in calls to the relational database.
- (4) How to solve the problem of termination for recursive predicates in the compiled approaches.

## REFERENCES

1. Kowalski, R., "Logic as a Database Language," Preprint, Imperial College, London, July 1981.
2. Gallaire, H., "Impacts of Logic on Databases," Proc. of 7th International Conf. on Very Large Databases, France, 1981.
3. Lloyd, J.W., "An Introduction to Deductive Database Systems," The Australian Computer Journal, Vol. 15, No. 2, May 1983, pp. 52-57.
4. Gallaire, H., "Logic Databases vs. Deductive Databases," Proc. of Logic Programming Workshop '83, Portugal, July 1983, pp. 608-622.
5. Chakravarthy, U., J. Minker and D. Tran, "Interfacing Predicate Logic Languages and Relational Databases," Proc. of 1st International Logic Programming Conference, France, 1982, pp. 91-98.
6. Minker J., "An Experimental Relational Database System Based on Logic," in Logic and Data Bases (Plenum Press, NY, 1978), pp. 107-148.
7. Kellogg, C. and L. Travis, "Reasoning with Data in a Deductively Augmented Data Management System," in Advances in Database Theory, Vol. I (Plenum Press, NY, 1980), pp. 261-295.
8. Kellogg, C. and L. Travis, "Deductive Planning and Pathfinding for Relational Databases," in Logic and Data Bases (Plenum Press, NY, 1978), pp. 279-200.
9. Chang, C.L., "DEDUCE2: Further Investigations of Deduction in Relational Data Bases," in Logic and Data Bases (Plenum Press, NY, 1978), pp. 201-236.
10. Chang, C.L., "On Evaluation of Queries Containing Derived Relations in a Relational Data Base," in Advances in Database Theory, Vol. I (Plenum Press, NY (1980), pp. 235-260.
11. Warren, D.H.D., "Efficient Processing of Interactive Relational Database Queries Expressed in Logic," Proc. of 7th International Conf. on Very Large Databases, France, 1981, pp. 272-281.

12. Grant, J. and J. Minker, "Optimization in Deductive and Conventional Relational Database Systems," in Advances in Database Theory, Vol. I (Plenum Press, NY, 1980), pp. 195-234.
13. Reiter, R., "Deductive Question-Answering on Relational Databases," in Logic and Data Bases (Plenum Press, NY, 1978), pp. 143-178.
14. Henschen, L., and S. Naqvi, "On Compiling Queries in Recursive First-Order Databases," Journal of the ACM, Vol. 31, No. 1, Jan. 1984, pp. 47-85.
15. Klahr, P., "Planning Techniques for Rule Selection in Deductive Question-Answering," in Pattern-Directed Inference Systems (Academic Press, NY, 1978), pp. 223-239.
16. Kahn, K.M. and M. Carlsson, LM-Prolog User Manual, Release 1.0, UPMAIL Report, Uppsala University, Sweden, Sept. 1983.
17. Britton Lee, Inc., IDM Software Reference Manual, Version 1.4, Los Gatos, California, Jan. 1983.



**END**

**FILMED**

**7-85**

**DTIC**